Improving fertility care

The role of guidelines, quality indicators and patients

Selma Mourad

The research described in this thesis was part of the 'SPRING'-study, which was supported by a grant from The Netherlands Organization for Health Research and Development (ZonMw), grant no. 945-14-116, and registered as trial no. NCT00119925 at www.ClinicalTrials.gov .

All studies were conducted at the Scientific Institute for Quality in Healthcare (IQ healthcare) and the department of Obstetrics and Gynaecology, both situated at the Radboud University Nijmegen Medical Centre. Both are part of the Nijmegen Centre for Evidence Based Practice (NCEBP), one of the approved research institutes of the Radboud University Nijmegen.

For reasons of consistency, terminology may be changed throughout this thesis when compared to the orginal publications.

Financial support for printing of this thesis was kindly provided by:
Scientific Institute for Quality of Healthcare (IQ healthcare), Ferring BV,
Schering-Plough Nederland , Goodlife BV, Merck Serono, Medical Dynamics

Selma Mourad, Nijmegen 2010

ISBN: 978-90-76316-80-2

Printing: Ipskamp Drukkers, Nijmegen

Layout: Jolanda van Haren & Selma Mourad

Cover design: Lonneke van Druten & Selma Mourad

Improving fertility care

The role of guidelines, quality indicators and patients

een wetenschappelijke proeve
op het gebied van de Medische Wetenschappen

Proefschrift

Ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. mr. S.C.J.J. Kortmann,
volgens besluit van het college van decanen
in het openbaar te verdedigen op woensdag 21 april 2010
om 15.30 uur precies

door

Salma Maria Mourad

geboren te Leiden op 1 maart 1980

Promotores:	Prof. dr. R.P.T.M. Grol Prof. dr. J.A.M. Kremer
Copromotores:	Mw. dr. R.P.M.G. Hermens Mw. dr. W.L.D.M. Nelen
Manuscriptcommissie:	Prof. dr. J. van der Velden Prof. dr. H.P.H. Kremer (UMCG) Prof. dr. J.M.M. van Lith (LUMC)
Paranimfen:	Mw. E.C. Haagen Mw. J.V. Stam

Improving fertility care

The role of guidelines, quality indicators and patients

كل عقدة لها حلاّل

'Every knot has someone to undo it'

Contents

Chapter 1

General Introduction

Introduction

Infertility should be considered one of the major health problems of the 21st century. It is commonly defined as 'any form of reduced fertility with prolonged time of unwanted non-conception'.[1] The worldwide prevalence of infertility is estimated to range from 4 to 30%, affecting approximately 80 million couples around the globe.[2,3] In industrialized countries, the lifetime prevalence of infertility is described to range from 16-26%, thus affecting almost 1 out of every 4 couples.[4]

This thesis focuses on studying the quality of current fertility care in the Netherlands and the implementation of the guidelines developed in the Dutch guideline programme on fertility care. This introductory chapter will start with a description of the prevalence of infertility in the Netherlands and the treatment options that are currently available. The impact on the Dutch society and healthcare system is discussed. Next, the quality of care and the potential role of clinical practice guidelines and quality indicators in the assessment and improvement of current care, will be addressed. The chapter concludes with a set of research questions that led to the studies performed within this thesis and a brief outline of the thesis chapters.

Infertility in the Netherlands

In the Netherlands, the number of new cases of infertility in general practice is estimated as two per 1000 couples per year.[5] In total, this concerns a referral to secondary or tertiary care for about 15 percent of couples of reproductive age.[5,6] After one year of unwanted non-conception, an initial assessment of fertility is performed, roughly consisting of three parts: an assessment of ovulation, an assessment of tubal patency and a semen-analysis.

The cause of the fertility problem is attributable to the male partner in roughly 30% of cases and to the female partner in another 30% of cases. A combination of both partners accounts for another 30% of identified causes, whereas it remains idiopathic for about 10% of couples. Depending on the identified cause of the fertility problem, the woman's age and duration of infertility, either an expectant regimen will be agreed upon, or a treatment plan will be formulated. Treatment can encompass fertility enhancing surgery or assisted reproductive technologies (ART) like ovulation induction (OI), Intra Uterine Insemination (IUI) with or without ovarian stimulation, In Vitro Fertilization (IVF) or Intra Cytoplasmatic Sperm Injection (ICSI) with either ejaculated or surgically retrieved semen. If these methods fail or are not suitable to the couple, alternative ways of conceiving or parenthood could be discussed, for instance, the use of donor gametes or opting for a surrogate parent or adoption.

In the Netherlands, fertility care is coordinated and mainly provided by gynaecologists and fertility clinicians, although general practitioners and urologists can, prior to referral to a gynaecologist, already start an initial assessment of fertility. Gynaecology departments in all clinics deliver primary fertility care, covering the initial assessment of fertility, surgery, OI and IUI. Performing a complete IVF or ICSI treatment (i.e. secondary fertility care) is limited to licensed clinics with highly specialized and accredited laboratories, of which there are 13 in the Netherlands. However, these 13 licensed clinics have affiliated regional clinics where the stimulation phase of the IVF or ICSI treatment can take place, respectively up to ovum pick-up (i.e. in 'satellite clinics') or including ovum pick-up (i.e. in 'transport clinics'). Thanks to these affiliations, a total

of 42 clinics in the Netherlands can offer their patients an IVF or ICSI treatment. The licensed clinics thus occupy a central position in regional fertility care. Tertiary fertility care comprises techniques like ICSI in combination with surgically retrieved epididymal or testicular semen; this care still remains restricted to a few of the licensed clinics and is partially performed in a research setting.

When compared to other countries, the accessibility to ART in the Netherlands is fortunately rather good. The basic Dutch health insurance covers the initial assessment of fertility, and the medical and medication costs of OI, IUI and three cycles of IVF or ICSI. As a result of this rather good accessibility to clinical care, 1 out of every 39 Dutch children born in 2007 was conceived by means of an IVF- or ICSI-treatment.[7] Taking all other forms of ART into account, this number might very well be even larger. However, infertility and its treatment is not to be taken lightly; it is associated with a large physical, social and psychological burden for the patients involved.[8-13] Apart from this impact on patients' lives, infertility claims an important position in health care due to the extensive use and costs of healthcare resources for diagnostics and treatment as well as for emotional and psychological support.

Quality of health care and the role of clinical practice guidelines

Considering this vast extent of psycho-social burden and resources needed, it is of great importance that delivered infertility care is of high quality. The Institute of Medicine defines quality of health care as *"The degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge"*. As simple and obvious as this definition may sound, defining and operationalizing the concept 'quality' for different fields of health care or specific clinical conditions, is refractory. In 1992, the first educational paper on evidence-based medicine (EBM) was published[14], which initiated a shift from intuitive and opinion led medicine towards EBM. This resulted in a plethora of papers on EBM, and the evidence base of existing treatment methods became subject to research. Subsequently, a multitude of large randomized controlled trials (RCT's) and meta-analyses were performed, which are still considered the 'gold standard' to compose a solid evidence base for clinical problems. It is therefore not surprising that institutions like the Cochrane Collaboration were founded[15], who made it their prime purpose to review existing evidence and systematically gather and disseminate this knowledge. To date, it is widely accepted that health care is only considered of high quality when it is, amongst other dimensions, based on the best available scientific knowledge.[16] However, medical literature from several countries suggests that approximately 30–40% of patients still do not receive care based on the best available scientific evidence. The quality of health care is moreover reported to vary between different settings[17] and an estimated 20–25% of provided health care is even considered unnecessary or even potentially harmful.[18-20] In the case of infertility, this could mean the use of unnecessary and expensive diagnostic tests and assisted reproductive technologies (ART) or the realization of complicated high-order pregnancies. This may have substantial physical and psychological consequences for the patients involved.

In an effort to provide clinicians easily with information regarding optimal health care, clinical practice guidelines (CPG's) were introduced to bridge the gap between best

available evidence and the daily care delivered in the consulting room. These CPG's, based on solid research and valuable clinical experiences, are increasingly seen as one of the crucial tools for improving quality of care.[20,21] They may help to increase the efficiency of care and reduce variation in performance between different professionals and hospital settings.

For reproductive medicine, several large professional organizations have developed clinical guidelines for fertility care in the past few years. Among those are the World Health Organization (WHO)[22], the European Society for Human Reproduction and Embryology (ESHRE)[23], the National Institute of Clinical Excellence (NICE)[24], the American Society of Reproductive Medicine (ASRM)[25], as well as several smaller country-specific fertility societies, like in Denmark and Austria[26], and even third-party initiatives like the Institute for Clinical Systems Improvement (ICSI)[27]. In the Netherlands, the Dutch Society of Obstetrics and Gynaecology (NVOG) was one of the first medical specialist organizations to issue a broad guideline programme. To date, the NVOG guideline programme contains a total of 72 evidence- and consensus-based guidelines, all developed by a systematic NVOG-procedure (www.nvog.nl). For infertility, nine guidelines were developed and afterwards authorized in the general assembly. They comprise the entire spectrum of comprehensive fertility care, from the initial assessment of fertility and diagnostics to treatment possibilities and complications. In addition, the Dutch Ministry of Health, Welfare and Sport published a model protocol of the Dutch Embryo Act (http://www.vsop.nl/pdf/embryowet.pdf), which incorporates a diversity of clinical recommendations concerning IVF/ICSI treatment, and which is therefore considered a clinical guideline as well. The studies in this thesis will mostly focus on a set of national infertility guidelines in the Netherlands, which is shown in table I.

Table I; The Dutch infertility guidelines used for the studies in this thesis

Guideline	Issued by	Year of publication
Initial Assessment of Fertility	NVOG	2005
Anovulation and Childwish	NVOG	2004
Male Infertility	NVOG	1998
Tubal Pathology	NVOG	2005
Endometriosis	NVOG	2001[a]
Premature Ovarian Failure (POF)	NVOG	2001
Intra Uterine Insemination (IUI)	NVOG	2000
Indications for In Vitro Fertilization treatment	NVOG	1998
Ovarian Hyper Stimulation Syndrome (OHSS)	NVOG	1998[b]
Embryo Act	VWS	2002

[a] revised in 2006; [b] revised in 2008; both revisions not considered for the studies within this thesis
NVOG=Dutch Society of Obstetrics and Gynaecology
VWS= Ministry of Health, Welfare and Sports

Assessing the quality of healthcare by quality indicators

However, development and dissemination of guidelines does not warrant their implementation in daily practice.[28] Former experience from NICE[29] and a multicentre study in the Netherlands directed at the use of the NVOG guideline on IUI care, has learned that current infertility guideline adherence is not optimal (unpublished data).

It can be assumed that this is also true for the remaining Dutch infertility guidelines, as they concern not only largely the same patient group, but are also used by exactly the same professionals.

A crucial first step towards improving guideline adherence is to gain insight into the application of clinical guidelines in daily practice, i.e. the current care. To estimate such quality of current care, valid quality indicators are necessary instruments.[30] A Quality indicator can be defined as *"a measurable element of practice performance for which there is evidence or consensus that it can be used to assess the quality of care provided, and hence change the quality of care provided".*[31]. Such quality indicators should preferably be developed by means of a systematic approach.[32] In general, three types of quality indicators can be distinguished, referring respectively to process and structure of medical care and the outcome of delivered care.[33] Indicators are constructed of a numerator (the cases in which a recommendation is followed) and denominator (all cases in which the recommendation is applicable). For fertility care, an example of each of these three types is given in table II.

Table II; Examples of three types of quality indicators for fertility care

Indicator type	Example of indicator construct
Process	Number of transfers in which no more than two embryos are transferred / Number of embryotransfers performed
Structure	Number of departments performing IUI who evaluate results annually / Number of departments performing IUI
Outcome	Number of started cycle of IVF/ICSI treatment that result in a life birth / Number of started cycles of IVF/ICSI treatment

Although the increasing need for evidence-based quality indicators in reproductive health care is recognized internationally, the search for the most suitable clinical indicators is still ongoing.[34-36] The sparse initiatives taken in this direction are not satisfactory yet; they concern either public health oriented initiatives[37-39] or lack a systematic development methodology.[24] Therefore, completing existing infertility guidelines with systematically developed quality indicators, would be a valuable step forward in quality of care improvement.

Improving the quality of health care by guideline implementation

Nonetheless, even if such systematically developed quality indicators are available, improving guideline implementation is a complex and challenging task.

Grol and others propose a stepwise cyclical model to implement change in clinical practice.[28,40] This model starts off with an assessment of current care, for example by means of quality indicators. Based on this, a change proposal can be made, which has to be based on available evidence or consensus. Then, barriers and facilitators to change should be identified, after which a tailored intervention can be developed. This

intervention can then be disseminated and implemented. After the implementation period, the intervention has to be evaluated, if possible in a RCT or otherwise in a large observational study. If the targets set beforehand are not achieved, any of the preceding steps can be modified and followed anew. If the targets are indeed achieved, it still remains important to keep monitoring current care, as the sustainability of change is often questioned.

Regarding the implementation of guidelines in general, literature from several fields of healthcare discloses that there is no generic 'magic bullet'.[40-42] Implementation strategies that appear to be successful and cost-effective in one clinical setting can easily fail in another context. Due to the lack of methodologically strong studies on guideline implementation strategies, the available evidence base continues to be imperfect.[41] Moreover, the shortage of publications on process evaluations concerning tested strategies, makes it hard to gain further insight into the 'black box' of events during implementation[43]; it thus remains unclear why some interventions do and some do not bring about the desired effects.

Multifaceted interventions were previously thought to be more effective than single interventions[44-46], but recent literature questions this anew[41]. However, a prospective identification of specific barriers to change and tailoring interventions to those barriers seems, though sometimes costly and time-consuming, to be still the most suitable way to develop an implementation strategy.[40,41,47,48] The most frequently tested and studied interventions to implement guidelines encompass audit and feedback on current practice, dissemination of educational materials, reminders and the organization of educational meetings or outreach visits. However, in general, mainly organizational aspects and individual professionals are targeted in these implementation interventions. It is quite surprising that the role of patients in guideline implementation is still rather unexplored.[40,49,50] For reproductive medicine or fertility care in particular, implementation research is still very rarely performed, leading to a lack of experiences and empirical data.

Patients as allies in guideline implementation

The concept of patient-centeredness and active patient participation in health care increasingly gains importance.[16] This is, for instance, reflected in the development of numerous patient surveys on preferences, experiences and satisfaction with care, the development of decision aids for diverse clinical problems and the increasing uptake of patients as partners in professional guideline development. All these innovations aim to empower patients, as there is evidence that active patient participation is associated with better medical care and improved health outcomes.[49,51-53] Although reviews of patient-oriented interventions to change clinical practice so far show mixed effects[54], there might be a crucial role set aside for patients within guideline implementation. As some of the frequently described barriers to guideline implementation in reproductive medicine concern the patients themselves[55,56], implementation of infertility guidelines might benefit from lifting these barriers by involving patients in the implementation process. We hypothesize patients could par excellence be the allies we need for successful implementation of the Dutch fertility guidelines, as fertility patients are, compared to other patient groups, generally young and critical towards their care providers. Moreover, Dutch fertility patients are united

in the large and very active patient association 'Freya', which not only constitutes a source of information but also advocates patients rights and viewpoints.

In conclusion, there exists a considerable discrepancy between the worldwide development and dissemination of clinical practice guidelines for fertility care on the one hand, and the lack of knowledge on implementation of such guidelines on the other hand. It is exactly this discrepancy that imposes the necessity to perform more research in this field.

Study objectives

This thesis studies the quality of current fertility care and the improvement of the implementation of infertility guidelines in the Netherlands, while focusing on the role of patients. The research concerning the implementation of the Dutch infertility guideline programme, will follow the implementation model of Grol et al.[40].

The research questions of the separate studies were as follows:

1. What can be regarded as a valid set of quality indicators for comprehensive fertility care? (chapter 2,3,4)

2. What is the quality of current fertility care in the Netherlands, and how is this experienced by patients? (chapter 4,5,6)

3. Does a combined professional and patient oriented intervention strategy prove more effective for the implementation of a infertility guideline programme than an exclusively professional oriented strategy? (chapter 7)

4. What are successful elements of the intervention strategies used? (chapter 7)

Outline of this thesis

At first sight, infertility may seem one of few fields in health care for which defining quality indicators of good clinical care is a really easy task; don't fertility clinicians all strive to enable their patients to conceive a healthy newborn? However, taking only such an evident outcome measure into account as a predictor of good care, does not give us any starting-points for quality improvement programmes. Chapter 2 debates this issue and pleads for the development of evidence-based quality indicators for reproductive medicine. Chapter 3 subsequently describes the systematic development of a set of quality indicators for fertility care, which are based on international literature, the 9 Dutch NVOG-guidelines and the Model protocol of the Dutch Embryo Act. In chapter 4, this set of quality indicators are submitted to a practice test in 16 Dutch fertility clinics, to assess several quality criteria per indicator and thus investigate their value for monitoring and improvement of clinical performance. A second aim, conditional on sufficient quality of the instrument, will be to assess the variation in current fertility care in a large sample of clinics. The objective of the study described in chapter 5 is to assess another aspect of current care; information provision to infertile couples. Adherence to guideline-recommendations on information provision will be examined by means of a patient survey. We will also assess patient satisfaction with this current practice and analyse to which extent variation in adherence is related to certain patient or clinic characteristics. The study described in chapter 6 is performed to gain more insight into the factors associated with positive patient experiences and satisfaction with fertility care. Furthermore, based on the results of the practice test described in chapter 4, two intervention strategies are developed to improve the implementation of the Dutch infertility guideline programme; a minimal strategy aimed at professionals and a maximal multi-faceted strategy aimed at both professionals and patients. In chapter 7, we will describe the large cluster randomized trial in which these two implementation strategies are tested. It also reports on the process evaluation of, and the professionals' and patients' experiences with, these implementation strategies.

This thesis concludes with a general discussion in chapter 8, in which we reflect on the findings from the previous chapters while discussing the results and methodology. We will also present the main conclusions of the studies presented in this thesis, and discuss the implications for future practice as well as future research.

References

1. Gnoth C, Godehardt E, Frank-Herrmann P, Friol K, Tigges J, Freundl G. Definition and prevalence of infertility and infertility. Hum Reprod 2005;20:1144-7.
2. Vayena E, Rowe PJ, Griffin P. Current practices and controversies in assisted reproduction: report of a WHO meeting. Geneva: World Health Organisation. 2002
3. Nachtigall RD. International disparities in access to infertility services. Fertil Steril 2006;85:871-5.
4. Schmidt L, Holstein BE, Boivin J, Sangren H, Tjornhoj-Thomsen T, Blaabjerg J et al. Patients' attitudes to medical and psychosocial aspects of care in fertility clinics: findings from the Copenhagen Multi-centre Psychosocial Infertility (COMPI) Research Programme. Hum Reprod 2003;18:628-37.
5. Beurskens MP, Maas JW, Evers JL. [Infertility in South Limburg: calculation of incidence and appeal for specialist care]. Ned Tijdschr Geneeskd 1995;139:235-8.
6. Centraal Bureau voor Statistiek (CBS). Maandstatistiek van de Bevolking (de Jong A, Steenhof L). 2000
7. Kremer JAM. Stichting Landelijke Infertiliteit Registratie: Landelijke IVF-cijfers 1996-2007 (www.nvog.nl). 2008
8. Greil AL. Infertility and psychological distress: a critical review of the literature. Soc Sci Med 1997;45:1679-704.
9. Schmidt L. Psychosocial burden of infertility and assisted reproduction. Lancet 2006;367:379-80.
10. Verhaak CM, Smeenk JM, van MA, Kremer JA, Kraaimaat FW. A longitudinal, prospective study on emotional adjustment before, during and after consecutive fertility treatment cycles. Hum Reprod 2005;20:2253-60.
11. Kallen B, Finnstrom O, Nygren KG, Otterblad OP, Wennerholm UB. In vitro fertilisation in Sweden: obstetric characteristics, maternal morbidity and mortality. BJOG 2005;112:1529-35.
12. Ombelet W, Martens G, De SP, Gerris J, Bosmans E, Ruyssinck G et al. Perinatal outcome of 12,021 singleton and 3108 twin births after non-IVF-assisted reproduction: a cohort study. Hum Reprod 2006;21:1025-32.
13. Schieve LA, Cohen B, Nannini A, Ferre C, Reynolds MA, Zhang Z et al. A population-based study of maternal and perinatal outcomes associated with assisted reproductive technology in Massachusetts. Matern Child Health J 2007;11:517-25.
14. Evidence-based medicine working group. Evidence-based medicine. A new approach to teaching the practice of medicine. JAMA 1992;268:2420-5.
15. Cochrane Collaboration. www.cochrane.org. 2009
16. Institute of Medicine. Crossing the quality chasm: a new health system for the twenty-first century. 2001
17. Seddon ME, Marshall MN, Campbell SM, Roland MO. Systematic review of studies of quality of clinical care in general practice in the UK, Australia and New Zealand. Qual Health Care 2001;10:152-8.
18. McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, DeCristofaro A et al. The quality of health care delivered to adults in the United States. N Engl J Med 2003;348:2635-45.
19. Schuster MA, McGlynn EA, Brook RH. How good is the quality of health care in the United States? Milbank Q 1998;76:517-63, 509.
20. Grol R. Successes and failures in the implementation of evidence-based guidelines for clinical practice. Med Care 2001;39:II46-II54.
21. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. BMJ 1999;318:527-30.
22. Rowe PJ, Hargreave TB, Mellows HJ and Comhaire FH (1993) WHO Manual for the Standardized Investigation and Diagnosis of the Infertile Couple. Cambridge University Press, Cambridge.
23. ESHRE. www.eshre.com. 2009
24. NICE. Fertility: assessment and treatment for people with fertility problems (http:// www.nice.org.uk/nicemedia/pdf/CG011niceguideline.pdf). 2004
25. ASRM. http:// www.asrm.org. 2009
26. Haagen EC, Hermens RP, Nelen WL, Braat DD, Grol RP, Kremer JA. Infertility guidelines in Europe: the quantity and quality of intrauterine insemination guidelines. Hum Reprod 2006;21:2103-9.
27. ICSI. the Institute for Clinical System Improvement: The diagnosis and management of infertility

(http://www.icsi.org/infertility/diagnosis_and_management_of_infertility_2301.html). 2004
28. Grol R. Personal paper. Beliefs and evidence in changing clinical practice. BMJ 1997;315:418-21.
29. NICE. A survey measuring the impact of NICE Guideline 11:Fertility – assessment and treatment for people with fertility problems. Survey Two: Fertility Experts. 2005
30. Grol R, Baker R, Moss F. Quality improvement research: understanding the science of change in health care. Qual Saf Health Care 2002;11:110-1.
31. Lawrence M, Olesen F, et al. Indicators of quality in health care. Eur J Gen Pract 1997;3:103-8.
32. Campbell SM, Braspenning J, Hutchinson A, Marshall MN. Research methods used in developing and applying quality indicators in primary care. BMJ 2003;326:816-9.
33. Donabedian A. The quality of care. How can it be assessed? JAMA 1988;260:1743-8.
34. Jahn A, Bloemenkamp KW, Hannaford P, Olsen J, da Silva MO, Temmerman M. Monitoring reproductive health in Europe--what are the best indicators of reproductive health? Hum Reprod 2006;21:2199-200.
35. Min JK, Breheny SA, MacLachlan V, Healy DL. What is the most relevant standard of success in assisted reproduction? The singleton, term gestation, live birth rate per cycle initiated: the BESST endpoint for assisted reproduction. Hum Reprod 2004;19:3-7.
36. Nelen WL, Hermens RP, Mourad SM, Haagen EC, Grol RP, Kremer JA. Monitoring reproductive health in Europe: what are the best indicators of reproductive health? A need for evidence-based quality indicators of reproductive health care. Hum Reprod 2007;22:916-8.
37. WHO. Monitoring reproductive health: selecting a short list of national and global indicators. 1997
38. Temmerman M, Foster LB, Hannaford P, Cattaneo A, Olsen J, Bloemenkamp KW et al. Reproductive health indicators in the European Union: The REPROSTAT project. Eur J Obstet Gynecol Reprod Biol 2006;126:3-10.
39. Ford J, Nassar N, Sullivan EA, Chambers G, Lancaster P. Reproductive health indicators Australia. 2003
40. Grol R, Wensing M, Eccles M. Improving Patient Care; the implementation of change in clinical practice. London: Elsevier Butterworth Heinemann, 2005.
41. Grimshaw J, Eccles M, Thomas R, Maclennan G, Ramsay C, Fraser C, et al. Toward evidence-based quality improvement. Evidence (and its limitations) of the effectiveness of guideline dissemination and implementation strategies 1966-1998. J Gen Intern Med 2006;21(Suppl 2):S14-S20.
42. Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. Lancet 2003;362:1225-30.
43. Hulscher ME, Laurant MG, Grol RP. Process evaluation on quality improvement interventions. Qual Saf Health Care 2003;12:40-6.
44. Grimshaw J, Eccles M, Tetroe J. Implementing clinical guidelines: current evidence and future implications. J Contin Educ Health Prof 2004;24 Suppl 1:S31-S37.
45. Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing physician performance. A systematic review of the effect of continuing medical education strategies. JAMA 1995;274:700-5.
46. Wensing M, van der WT, Grol R. Implementing guidelines and innovations in general practice: which interventions are effective? Br J Gen Pract 1998;48:991-7.
47. Shaw B, Cheater F, Baker R, Gillies C, Hearnshaw H, Flottorp S et al. Tailored interventions to overcome identified barriers to change: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2005CD005470.
48. Chaillet N, Dube E, Dugas M, Audibert F, Tourigny C, Fraser WD et al. Evidence-based strategies for implementing guidelines in obstetrics: a systematic review. Obstet Gynecol 2006;108:1234-45.
49. Wetzels R, Harmsen M, van WC, Grol R, Wensing M. Interventions for improving older patients' involvement in primary care episodes. Cochrane Database Syst Rev 2007CD004273.
50. Griffin SJ, Kinmonth AL, Veltman MW, Gillard S, Grant J, Stewart M. Effect on health-related outcomes of interventions to alter the interaction between patients and practitioners: a systematic review of trials. Ann Fam Med 2004;2:595-608.
51. Kaplan SH, Greenfield S, Ware JE, Jr. Assessing the effects of physician-patient interactions on the outcomes of chronic disease. Med Care 1989;27:S110-S127.
52. Atkin PA, Stringer RS, Duffy JB, Elion C, Ferraris CS, Misrachi SR et al. The influence of information provided by patients on the accuracy of medication records. Med J Aust 1998;169:85-8.

53. Anderson RM, Funnell MM, Butler PM, Arnold MS, Fitzgerald JT, Feste CC. Patient empowerment. Results of a randomized controlled trial. Diabetes Care 1995;18:943-9.
54. Grol R. Improving the quality of medical care: building bridges among professional pride, payer profit, and patient satisfaction. JAMA 2001;286:2578-85.
55. Haagen EC, Nelen WL, Hermens RP, Braat DD, Grol RP, Kremer JA. Barriers to physician adherence to a infertility guideline. Hum Reprod 2005;20:3301-6.
56. Chaillet N, Dube E, Dugas M, Francoeur D, Dube J, Gagnon S, et al. Identifying barriers and facilitators towards implementing guidelines to reduce caesarean section rates in Quebec. Bull World Health Organ 2007;85:791-7.

Chapter 2

Monitoring reproductive health in Europe: what are the best indicators of reproductive health? A need for evidence-based quality indicators of reproductive health care

W.L.D.M. Nelen
R.P.M.G. Hermens
S.M. Mourad
E.C. Haagen
R.P.T.M. Grol
J.A.M. Kremer

Human Reproduction 2007; 22(4): 916–918.

Abstract

Monitoring reproductive health by the Reprostat indicators in Europe will facilitate the transparency of reproductive health as well as comparisons over time and between countries. However, for the monitoring and improvement of reproductive health care, we suggest the systematic development of evidence-based quality indicators, especially process and structure indicators.

Reprostat

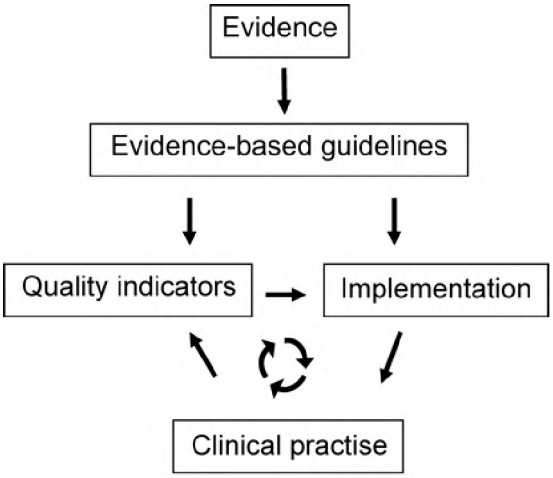
In accordance with Jahn et al.[1], we agree that monitoring reproductive health in Europe by indicators facilitates the transparency of reproductive health as well as comparisons over time and between countries. Moreover, insight into actual reproductive health may give rise to initiatives to improve reproductive health. Therefore, we underwrite the importance of the development of an indicator set for reproductive health by 'Reprostat' (http://ec.europa.eu/health/ph_projects/2001/monitoring/fp_monitoring_2001_a1_frep_02_en.pdf). However, the question is: what kind of indicators do we need in reproductive health? The Reprostat set contains mainly indicators for reproductive health and not for reproductive health care that we deliver in our clinics. Only a few of these indicators can be used as proxy measures for reproductive health care, such as Reprostat indicator 11: 'the proportion of deliveries associated with assisted reproductive technology (ART)'.

Monitoring reproductive health care

In our opinion, monitoring reproductive health care by quality indicators is also important, because it gives insight into the overall quality of delivered fertility care, it gives the opportunity to compare the delivered care with the recommended care in evidence-based guidelines and determination of substandard reproductive health care can easily guide improvement of this care. Therefore, to improve reproductive health care, a comprehensive set of clinical practice guidelines, valid quality indicators and effective strategies to implement the guidelines are needed.

Clinical practice guidelines provide clinicians easily accessible information regarding optimal reproductive health care. They are a tool to bridge the gap between evidence from the literature and the daily practice. However, clearly, the availability of evidence-based clinical guidelines by itself does not result in the delivery of optimal patient care.[2,3] Implementation of the key recommendations of these guidelines requires more than just their publication and dissemination. Well-developed and evaluated strategies are necessary to facilitate this implementation. Quality indicators, defined as 'measurable elements of practice performance for which there is evidence or consensus that they can be used to assess the quality of care' [4], are crucial in this field; they measure the application of guidelines in daily practice and provide ammunition for feedback and development of implementation strategies (Figure 1).

Figure 1; **Schematic representation of the different steps from clinical evidence to the implementation in clinical practice.**



Different types of indicators

Quality indicators can refer to processes and structures of medical care as well as to the outcome of delivered care.[5] In the case of reproductive health care, 'the proportion of professionals monitoring ovarian stimulation by ultrasound' and 'the proportion of professionals informing infertile women about their reduced fertility by smoking' are process indicators, 'the proportion of hospitals with a laboratory accreditation' is a structure indicator and 'the live birth rate per cycle' and 'the proportion of women having been screened for Chlamydia trachomatis before uterine instrumentation' are outcome indicators.

Reproductive health care has so far mainly been monitored by outcome indicators prompting a recent debate in this journal.[6] However, a disadvantage of using outcome indicators as a measurement of health care performance is the probability factor in health care. That means just the same medical treatment does not always have the same outcome and the other way around: an acceptable outcome does not have to be caused by a desirable treatment. Differences in outcome may also be due to case mix and the way of data collection.[7,8] In addition, one of the potential risks to report quality of medical care in terms of outcome measurements only is the refusal to treat patients with a poor prognosis. The solution for these problems is to rely more on process and structure than on outcome indicators.

Process indicators tend to be more sensitive and rarely confounded by other factors, if properly designed.[7,8] Moreover, process indicators can steer plans towards directed improvement activities.[9] Of particular value are process or structure indicators that have been demonstrated to have a link with reproductive health-care outcome. The better this association, the stronger the benefits of applying the quality indicators are in terms of, for example, improved ongoing pregnancy rates and reduced multiple pregnancy rates.

Process and structure indicators should be based ideally on evidence-based guidelines, literature and experts' opinions. Moreover, such a set of indicators should be

developed systematically and carefully, accepted by the target group and be sensitive to changes in the quality of care.[10,11]

A need for process and structure indicators

However, despite growing recognition of the importance of increased transparency and more rigorous monitoring of health-care performance to decrease the delivery of inappropriate medical care, there have been only a few process or structure indicators suggested for reproductive health care.

For example, the clinical guideline about fertility assessment and treatment for people with fertility problems developed by the National Institute for Clinical Excellence contains five measures that could be used.[12] Recently, we developed process and structure indicators for all national fertility guidelines of the Dutch Society for Obstetrics and Gynaecologists in two Dutch multicentred studies: (i) the Quality study on Intrauterine insemination in the Netherlands among Gynaecologists (KING study) and (ii) the study about fertility guidelines: Patient-Related Implementation in the Netherlands among Gynaecologists (SPRING study; description and trial protocol can be found at http:// clinicaltrials.gov/show/NCT00119925).

Role of ESHRE?

Last year, some new clinical guidelines were published by the European Society of Human Reproduction and Embryology (ESHRE) about, e.g. endometriosis, preimplantation genetic diagnosis and recurrent miscarriage.[13-15] Although these evidence-based guidelines were well developed, unfortunately they were not accompanied by a set of quality indicators.

Recently, after an exploration of all national guidelines about intrauterine insemination in Europe, we suggested the establishment of a central body for reproductive medicine in Europe with expertise in up-to-date guideline development methodology to improve the scientific validity and the international consensus and to reduce the inefficient use of resources.[16] We would argue for incorporating this central guideline development with a central indicator development to facilitate the evaluation of reproductive health care over time and between countries. The issue of whether or not ESHRE, e.g. under the auspices of the special interest group in Quality and Safety in ART, will act here as a central body should be discussed.

In conclusion, we welcome the Reprostat indicators for monitoring reproductive health in Europe and see it indeed as an important part of a quality control system. However, for the monitoring and improvement of reproductive health care, we suggest the central and systematic development of process and structure indicators that are based on evidence-based clinical guidelines.

References

1. Jahn A, Bloemenkamp KW, Hannaford P, Olsen J, da Silva MO, Temmerman M. Monitoring reproductive health in Europe—what are the best indicators of reproductive health? Hum Reprod 2006;21:2199–200.
2. Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. Lancet 1993;342:1317–22.
3. Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD, Thomson MA. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. The Cochrane Effective Practice and Organization of Care Review Group. BMJ 1998;317:465–8.
4. Donabedian A. The quality of care. How can it be assessed? JAMA 1988;260:1743–8.
5. Campbell SM, Roland MO, Buetow SA. Defining quality of care. Soc Sci Med 2000;51:1611–1625.
6. Min JK, Breheny SA, MacLachlan V, Healy DL. What is the most relevant standard of success in assisted reproduction? The singleton, term gestation, live birth rate per cycle initiated: the BESST endpoint for assisted reproduction. Hum Reprod 2004;19:3–7.
7. Eddy DM. Performance measurement: problems and solutions. Health Aff (Millwood) 1998;17:7–25.
8. Mant J. Process versus outcome indicators in the assessment of quality of health care. Int J Qual Health Care 2001;13:475–80.
9. Rubin HR, Pronovost P, Diette GB. The advantages and disadvantages of process-based measures of health care quality. Int J Qual Health Care 2001;13:469–74.
10. Rubin HR, Pronovost P, Diette GB. From a process of care to a measure: the development and testing of a quality indicator. Int J Qual Health Care 2001;13:489–96.
11. Campbell SM, Braspenning J, Hutchinson A, Marshall MN. Research methods used in developing and applying quality indicators in primary care. BMJ 2003;326:816–9.
12. National Institute for Clinical Excellence. Fertility Assessment and Treatment for People with Fertility Problems. Auditable Standards. London: RCOG Press, 2004: chapter 18, p. 138, www.nice.org.uk.
13. Kennedy S, Bergqvist A, Chapron C, D'Hooghe T, Dunselman G, Greb R, Hummelshoj L, Prentice A, Saridogan E. ESHRE guideline for the diagnosis and treatment of endometriosis. Hum Reprod 2005;10:2698–704.
14. Thornhill AR, Die-Smulders CE, Geraedts JP, Harper JC, Harton GL, Lavery SA, Moutou C, Robinson MD, Schmutzler AG, Scriven PN, et al. ESHRE PGD Consortium 'Best practice guidelines for clinical preimplantation genetic diagnosis (PGD) and preimplantation genetic screening (PGS)'. Hum Reprod 2005;20:35–48.
15. Jauniaux E, Farquharson RG, Christiansen OB, Exalto N. Evidence-based guidelines for the investigation and medical treatment of recurrent miscarriage. Hum Reprod 2006;21:2216–22.
16. Haagen EC, Hermens RP, Nelen WL, Braat DD, Grol RP, Kremer JA. Subfertility guidelines in Europe: the quantity and quality of intrauterine insemination guidelines. Hum Reprod 2006;21:2103–9.

Chapter 3

Guideline-based development of quality indicators for fertility care

S.M. Mourad
R.P.M.G. Hermens
W.L.D.M. Nelen
D.D.M. Braat
R.P.T.M. Grol
J.A.M. Kremer

Human Reproduction 2007;22(10):2665-72.

Abstract

Background: Internationally, several organizations have developed clinical guidelines for fertility care to supply patients with the best possible care. However, to improve the implementation of such guidelines, we first need to gain insight into the application of clinical guidelines in daily practice. Valid quality indicators are necessary to estimate actual guideline adherence. However, none of the existing fertility guideline programmes is accompanied by a satisfactory set of quality indicators. In this study, we develop a set of valid guideline-based quality indicators for fertility care.

Methods: A systematic RAND-modified Delphi method was used to develop a set of key recommendations based on 10 national Dutch fertility guidelines, international literature, and existing international indicators. Experts' opinions were used to appraise recommendations regarding specific criteria such as efficacy, level of health gain, applicability, and potential for care improvement.

Results: A representative set of 39 key recommendations was selected from 303 initial recommendations. The recommendations covered two structural and 37 procedural aspects, the latter encompassing 'indications for treatment', 'diagnostic procedures', 'treatment procedures', and 'patient information'.

Conclusions: This study describes the systematic, stepwise method used to develop 39 process and structure indicators that can be used to monitor fertility care.

Introduction

The worldwide prevalence of infertility is estimated to range from 4% to 30%, affecting approximately 80 million couples around the globe.[1,2] Infertility should therefore be considered one of the major health problems of the 21st century; it has great societal impact. Even if sufficient fertility services are available (as is the case in most western countries), they are not necessarily accessible for all couples nor is the quality necessarily satisfactory. The available diagnostic tests and assisted reproductive technologies (ART) are expensive and can have substantial physical and psychological consequences for the patients involved.[3,4] Pregnancies after ART are more often complicated, particularly due to higher multiple pregnancy rates.[5-7] Therefore, optimal organization of care should be strived for in order to supply patients with the best possible care and a minimum of complications at minimal costs. However, the question rises whether the current organization of fertility care serves this goal satisfactorily, and to what extent patients effectively receive such 'best care'.

To help physicians standardize fertility care and improve health outcomes, several large professional organizations have developed clinical guidelines for fertility care in the past few years. These include the National Institute of Clinical Excellence (NICE), the American Society of Reproductive Medicine (ASRM), the European Society for Human Reproduction and Embryology (ESHRE), as well as several country-specific fertility societies. In the Netherlands, the Dutch Society for Obstetrics and Gynaecology (NVOG) has issued nine national guidelines for fertility care. In general, clinical practice guidelines, preferably based on solid research and valuable clinical experience, are increasingly seen as one of the crucial tools for achieving high quality of care. [8] Guidelines can easily provide clinicians with information regarding optimal health care, and they aim at increasing the efficiency of care and reducing variation in performance between different professionals and hospitals. Regarding fertility care, adherence to guidelines can decrease unnecessary diagnostics and treatments, which reduces the number of complications and expenditures. However, publication and dissemination of guidelines is not automatically followed by a change or improvement of health care[9]; guidelines do not implement themselves[10]. From our own experience in a multicentre quality study of intrauterine insemination (IUI) in the Netherlands, we have learned that current guideline adherence is not optimal (unpublished data). The first step toward improving adherence is to gain insight into the application of clinical guidelines in daily practice, i.e. the actual care. To estimate such quality of actual care, valid quality indicators are necessary[11] because they are the missing link in connecting evidence from guidelines to practice. Three types of quality indicators can be distinguished referring to the process and structure of medical care and the outcome of delivered care. Such quality indicators should preferably be developed by means of a systematic method.[12] The increasing need for evidence-based quality indicators in reproductive health care is recognized internationally as well, but the search for the most suitable indicators is still ongoing.[13-15] Unfortunately, none of the existing fertility guidelines is accompanied by a satisfactory set of systematically developed quality indicators. In our current study, we therefore aimed to develop a set of valid quality indicators covering all aspects of fertility care.

Materials and methods

Setting

A systematic RAND-modified Delphi method[16,17], including independent expert ratings and repetitive feedback, was used to develop a set of key recommendations suitable for transcription into quality indicators. We based our study on international literature and the nine NVOG fertility guidelines: initial assessment of fertility, anovulation and child wish, male infertility, tubal pathology, endometriosis, premature ovarian failure (POF), intra uterine insemination (IUI), indications for in vitro fertilization (IVF) treatment and ovarian hyperstimulation syndrome (OHSS) (http://www.nvog-documenten.nl, in Dutch). These guidelines are consensus or evidence based and were systematically developed according to NVOG standards and issued between 1998 and 2005. In addition, the Dutch Ministry of Health, Welfare and Sport published a model protocol of the Dutch Embryo Act (http://www.vsop.nl/pdf/embryowet.pdf, in Dutch), which incorporates a diversity of clinical recommendations concerning IVF treatment, and which is considered a clinical guideline as well. The invited expert panels consisted of a representative diversity of clinical experts and guideline users[18], working at various types of fertility clinics ranging from small regional hospitals to tertiary university clinics. Among them were gynaecologists, fertility clinicians (medical doctors who specialize in assisted reproduction, but who do not have a postgraduate specialization in obstetrics and gynaecology), gynaecologists in training, and members of the original guideline development workgroups.

Procedure for indicator development

The procedure for quality indicator development consisted of six steps, which were completed consecutively as shown in Figure I: 1) literature search, 2) selection of recommendations, 3) written questionnaire, 4) consensus meeting, 5) critical evaluation, and 6) consultation with guideline developers:

1. Literature search

We searched Medline for existing indicators for reproductive health and specifically fertility care. The key words used were 'subfertility', 'infertility', and 'reproductive health care' combined with 'guidelines' or 'clinical guidelines' and 'quality indicators'. Furthermore, we searched the Internet for potential sources of quality indicators, e.g. governmental reports, reports from national and international fertility societies, and commercial initiatives. We paid specific attention to existing guidelines concerning reproductive health or infertility. For this purpose we used the Websites of the National Guideline Clearinghouse (www.guidelines.gov) and the Geneva Foundation for Medical Education and Research (www.gfmer.ch).

2. Selection of recommendations

Firstly, four authors (SM, RH, WN, and JK) separately selected all the recommendations from the nine NVOG guidelines and the Dutch Embryo Act. To take international consensus and solid evidence into account as well, we added all the recommendations with A-level evidence from the NICE fertility guideline that did not have an exact equivalent in the Dutch set of recommendations. The collected recommendations were edited in three different questionnaires for the consensus panel method in step

3. Questionnaire 1 covered 'endometriosis', 'anovulation', and 'IUI' (131 items); questionnaire 2, 'initial assessment of fertility', 'male infertility', and 'tubal pathology' (110 items); and questionnaire 3, 'indications for IVF', 'OHSS', 'POF' and the 'Embryo Act' (99 items).

3. Written questionnaire

Questionnaires were sent to 63 experts who were divided into three expert panels of 21 members each, and the panels appraised the questionnaires according to a RAND-modified Delphi method[12]. Each panel received the questionnaire corresponding to the area of interest of its members. Whenever available in the guidelines, international evidence levels were provided to facilitate the decision making. Experts were asked to score the key recommendations on a nine-point Likert scale ranging from 1 (hardly relevant) to 9 (extremely relevant), with respect to their impact on both 'health gain' ,i.e. health as 'physical, mental, and social well-being' [19], and 'overall efficacy', i.e. prevention of unnecessary medical treatment and promotion of cost-effectiveness. Moreover, the experts were asked to provide for each guideline a top-five ranking of recommendations they considered 'most important' and 'representative' to assess the quality of clinical performance. They were also offered an opportunity to provide comments and suggest additional items. First, we analysed the data from the returned questionnaires and calculated median scores for 'health gain' and 'overall efficacy' for each recommendation. The selected recommendations were rated as valid if they matched the criterion described by Campbell[20]: an overall panel median score of 8 or 9. To rate a recommendation as valid, there also had to be 'agreement' between the ratings of the independent panellists. Agreement was defined as 75% or more of ratings within a panel being in the lowest tertile (1, 2, 3) or the highest tertile (7, 8, 9), so that items that were mainly in the unequivocal tertile (4, 5, 6) were excluded.

Second, we scored the recommendations rated valid by means of the Campbell criteria by awarding them points according to the expert's top-five ranking. For each number-one ranking by an expert, we awarded a recommendation 5 points, for each number-two ranking, we awarded 4 points, and so on. This way, we created a list of scores reflecting the weight that experts assigned to each recommendation. For each guideline, the recommendations with a value of three points or more were listed top-item down. Next, a consensus meeting took place with these listings as feedback.

4. Consensus meeting

All participants in the questionnaire round were invited to a consensus meeting, where a discussion about the results of step 3 was initiated. The participants were divided into three panels, each assessing separate guidelines. Each panel was asked to discuss and reconsider the previous rankings and aim at a new top five; selected recommendations had to be 'applicable' for indicator development and had to have 'potential for improvement' as well. The panels were allowed to suggest additional recommendations. The discussions lasted approximately 2 hours and were chaired by four of the authors (SM, RH, WN, and JK), who are experienced in performance measurement. This resulted in a final consensus-based top five for each guideline, ready for consideration in step 5.

5. Critical evaluation
Four authors (SM, RH, WN, and JK) critically evaluated the top-five rankings from step 4. The recommendations were once more appraised with regard to both 'improvement potential' and 'applicability' after operationalization for the clinical setting. Some items were excluded or combined due to overlap between guidelines or pragmatic reasons concerning operationalization.

6. Consultation with guideline developers.
As a sixth and final step, we presented the final top-five rankings in a third feedback round to 12 gynaecologists who had participated in the development of these guidelines, and are therefore considered opinion leaders in the field.

Results

Step 1

The online search of Medline and the additional Internet search revealed that several national and international organizations issued specific guideline programmes in the field of reproductive health or infertility. Nevertheless, even widely distributed programmes, such as the ASRM and ESHRE programmes, are usually unaccompanied by quality indicators. However, programmes that are indeed accompanied by indicator sets[21-24] focus mainly on a public health perspective. They aim, for example, at monitoring prevalences of contraceptive use, Chlamydia infections, and induced abortions to describe current reproductive health care, but are clearly not specifically developed for the extensive domain of clinical outpatient fertility care. In contrast, the National Institute of Clinical Excellence (NICE) issued their Clinical Guideline (CG) 11 concerning infertility[25] , and it was accompanied by seven 'key priorities for implementation'. Because of the lack of other suitable and systematically developed infertility indicators for clinical care, we decided to include all 28 level-A-evidence recommendations from the CG11 for appraisal in our selection procedure (Figure I).

Step 2

Altogether, 275 recommendations extracted from the NVOG guidelines were added. This resulted in a total of 303 recommendations for appraisal by the experts.

Step 3

Initially, 63 questionnaires were sent to the experts, of which 47 (75%) were returned fully completed. The response rates were 71% (n=15) for questionnaire 1, 76% (n=16) for questionnaire 2, and 76% (n=16) for questionnaire 3. The 303 preselected recommendations were reduced to 139 items in this selection round, while 36 new items were suggested. This resulted in 175 items for appraisal during the consensus round.

Step 4

Eleven experts attended the consensus meeting: 3 guideline developers, 6 gynaecologists, 1 fertility clinician, and 1 gynaecologist in training. They discussed and prioritized the 175 items, resulting in a new selection of 53 recommendations.

Step 5

Critical evaluation of the 53 recommendations led to the exclusion of 6 items, and another 16 were combined to 8 new items. A semifinal set of 39 key recommendations was offered to the guideline developers for approval.

Step 6

All the guideline developers approved of the set of 39 key recommendations, which was thus finalized (Table II).

Final set

These 39 key recommendations were transcribed into 39 quality indicators concerning either structural or procedural domains of fertility care. Table I reflects for each guideline the selection of key recommendations within the various domains. Both structural (n=2) and procedural (n=37) aspects were selected. The latter encompass 'indications for ART' (n=7), 'diagnostic procedures' (n=6), 'treatment procedures' (n=20) and 'patient information' (n=4). In expert rounds 3 and 4, 8 of 16 patient information items (50%) were selected, as were 24 of 132 treatment procedure items (18%) and 11 of 91 diagnostic items (12%).

Furthermore, 2 of 5 structure items (40%) concerning the annual evaluation of IUI and the accreditation of IVF laboratories were selected (Table II). Only one of the 36 newly suggested items (regarding lifestyle advice) was included in the final set. Of the 28 level-A-evidence recommendations from the NICE guideline that were initially added in the first step, only 2 were included in the final set. Another 9 (non-level-A evidence) recommendations were common to both the NICE and the NVOG guidelines (Table II).

Figure I a stepwise RAND-modified Delphi method to develop quality indicators for fertility care

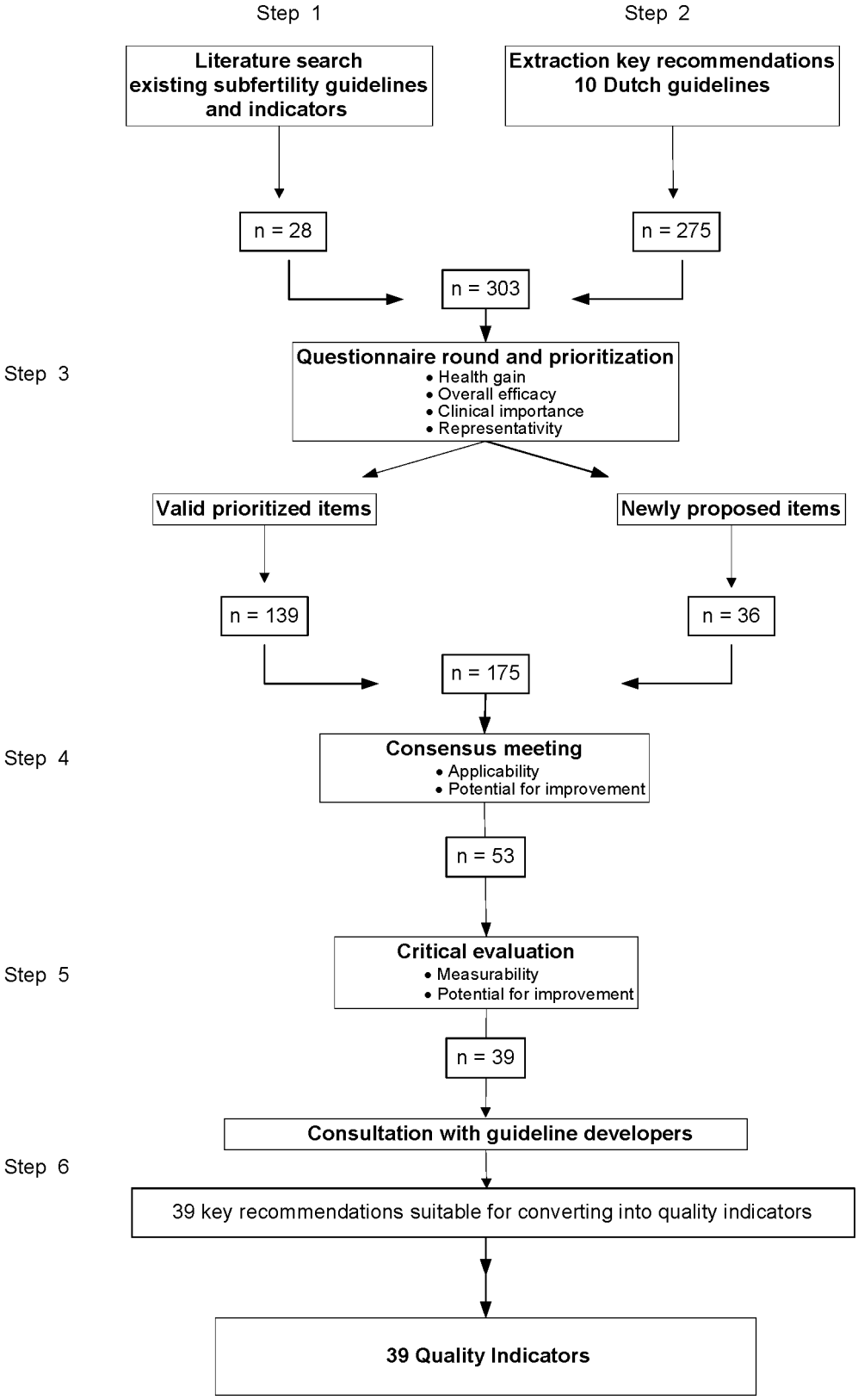


Table I: Different domains of key recommendations per guideline

Guidelines	OFO	Anovulation	Male factor	Tubal pathology	Endo-metriosis	POF	IUI	Indications for IVF	OHSS	Embryo Act	Domains after Steps 3 and 4	Domains in final set
Type of recommendation												
Process (total)												37
Diagnosis	4 (33)	(23)	3 (20)	1 (3)	3 (5)	(5)	(1)	–	(1)	–	11 (91)	6
(Contra-) Indications for ART	–	(3)	–	(2)	1 (1)	1 (1)	3 (5)	3 (19)	–	–	8 (31)	7
Treatment procedures	(5)	5 (37)	2 (8)	4 (16)	1 (4)	1 (2)	6[a] (24)	1[a]	3 (21)	1 (15)	24[b] (132)	20[b]
Patient information	1[c]	2 (4)	–	(1)	–	4 (5)	–	(1)	(4)	1 (1)	8[c] (16)	4[c]
Structure	–	–	–	–	–	–	1 (3)	–	(1)	1 (1)	2 (5)	2
Total selected recommendations	5 (38)	7 (67)	5 (28)	5 (22)	5 (10)	6 (13)	10 (33)	4 (20)	3 (27)	3 (17)	53 (275)	39
Total indicators formed	5	6	3	4	3	3	5	4	3	3		39

Every column per guideline shows the amount of key recommendations selected by the consensus-procedure; between parentheses is the total amount of recommendations eligible for selection.
OHSS, ovarian hyper stimulation syndrome; OFO, initial assessment of fertility; POF, premature ovarian failure.

[a] An item deriving from NICE guideline CG11.

[b] 2 items deriving from NICE guideline CG11.

[c] An item newly suggested by the expert panel.

Table II: Final set of key recommendations eligible for indicator transcription per guideline, indicator-type and origin

Guideline	Type of indicator	Guideline of origin
Initial assessment of fertility		
• The initial fertility assessment should result in both a diagnosis and a prognosis.	Process	NVOG
• The initial fertility assessment should consist of three parts: semen analysis, tuba-passage and cycle-analysis.	Process	NVOG
• Couple's history-taking should cover at least: age of both partners, duration of infertility, type of couple's infertility (primary or secondary).	Process	NVOG
• Woman's physical examination should include assessment of the body mass index.	Process	NVOG
• Life-style advice concerning bodyweight, smoking and alcohol and drug use should be part of the counselling regarding pregnancy-probabilities.	Process	Panel,NICE
Anovulation and childwish		
• Patients with overweight should, with regard to their fertility-treatment and overall health, be informed of the importance of weight reduction by means of life-style changes.	Process	NVOG,NICE
• The goal of ovulation-induction should be mono-ovulation.	Process	NVOG
• The ovarian response to hormonal stimulation should be performed by regular transvaginal ultrasound (frequency of 1–3 times/week).	Process	NVOG
• Women with WHO Group II anovulatory infertility should be given anti-oestrogen as first-choice medication for ovulation induction.	Process	NVOG
• Before starting ovulation induction treatment, the patient should be informed about specific side-effects of medication, the need for regular intensive follow-up during treatment, the increased risks of multiple pregnancy, OHSS syndrome and spontaneous abortion.	Process	NVOG
• In case of three or more follicles >16 mm and/or more than five follicles >12 mm during an ovulation-induction treatment cycle, the patient has to be informed that coitus is prohibited or contraception has to be used.	Process	NVOG
Male factor infertility		
• In case of normospermia (WHO-criteria), semen-analysis should not be repeated, and no additional andrological tests should be performed.	Process	NVOG
• In case of an abnormal semen-analysis (WHO-criteria), the physician should perform an andrological anamnesis, a physical examination and at least one extra semen-analysis.	Process	NVOG,NICE
• In case of an idiopathic oligoasthenoteratozoospermia, no hormones, vitamins or non-steroidal anti-inflammatory drugs should be prescribed to improve semen-quality.	Process	NVOG,NICE
Tubal pathology		
• In case of IVF because of inoperable tubapathology, salpingectomy shoud be performed when bilateral hydrosalpinges are visible by ultrasound.	Process	NVOG

Guideline	Type of indicator	Guideline of origin
• In case of tubal pathology, both diagnostic laparoscopy and complementary examination of the endosalpinx (hysterosalpingogram or salpingo-scopy) should be performed to decide in favour of tubal surgery.	Process	NVOG
• The indication for tubal surgery in a patient should be made by the future surgeon himself (i.e. videomaterial or by performing a diagnostic laparoscopy).	Process	NVOG
• Except for refertilization, the female age limit for tubal surgery should be 40 years.	Process	NVOG
Endometriosis		
• Prior to laparoscopy intended to diagnose endometriosis, at least an anamnesis should be performed for those with a suspect history, a vaginal and speculum examination.	Process	NVOG
• Women with peritoneal endometriosis and infertility should not be given hormonal treatment to improve fecundicity.	Process	NVOG,NICE
• In case of unexplained infertility and peritoneal endometriosis, IUI with ovarian stimulation should be given as primary treatment to improve fecundicity.	Process	NVOG
Premature ovarian failure (POF)		
• Patients with POF, should not be offered any treatment to pursue pregnancy, except for oocyte donation.	Process	NVOG,NICE
• In case POF is diagnosed, the following should be discussed with the patient: possibilities to receive psychological support, oocyte-donation, the probabilities of a spontaneous conception, the option of assessing karyotype (POF, Fragile X, Auto-immune diseases) and the (dis-)advantages of hormone replacement therapy	Process	NVOG
• In every patient with POF who chooses not to have hormonal supplementation therapy, bone-densitrometry should be performed.	Process	NVOG
Intra uterine insemination (IUI)		
• In case of unexplained infertility, stimulated IUI should not be offered, even though it is associated with higher pregnancy rates than unstimulated IUI, because it carries a risk of multiple pregnancy.	Process	NICE
• The diagnosis 'cervical factor' is an indication for IUI in the unstimulated cycle.	Process	NVOG
• In case of IUI in the stimulated cycle, ovarian response should be monitored by transvaginal ultrasound.	Process	NVOG
• Each department performing IUI should evaluate their results annually.	Structure	NVOG
• IUI should not be performed in case of more than three follicles .16 mm or more than five follicles .12 mm. In both cases, the use of contraceptives should be advised as well.	Process	NVOG
Indications for IVF/ICSI		
• In case of severe endometriosis and decreased tubal function (but no bilateral occlusion), there is an indication for IVF after 2 years of infertility.	Process	NVOG
• In case of male infertility with volume x concentration x motility < 1 x10^6/ml (before capacitation) there is a direct indication for ICSI-treatment.	Process	NVOG,NICE
• In case of unexplained infertility in a woman < 36 years, there is an indication for IVF after 3 years of infertility.	Process	NVOG,NICE
• Routine use of hCG for luteal support after IVF is not recommended.	Process	NICE

Guideline	Type of indicator	Guideline of origin
Ovarian hyper stimulation syndrome (OHSS)		
• In case of ovulation-induction with gonadotropins, no ovulatory HCG dose should be given and protected intercourse should be advised in case of the presence of three or more dominant follicles ≥ 18 mm or five or more follicles ≥ 15 mm and/or a serum–estradiol level > 3.0 nmol/l.	Process	NVOG
• In case of OHSS with haemoconcentration ≥ 45%, the patient should be admitted to the hospital.	Process	NVOG
• In case of severe OHSS and hospital admission, patients should be given thrombosis–prophylaxis (low molecular weight heparin).	Process	NVOG
Embryo act		
• IVF-laboratories should be accredited by the CCKL (Dutch Coordinating Committee for maintenance of Quality Standards in Laboratory research).	Structure	NVOG
• Per embryo transfer no more than two embryos should be transferred at the same time.	Process	NVOG,NICE
• During an intake prior to IVF-treatment, the following should be discussed:	Process	NVOG
◦ The risks of hyperstimulation, poor response and accompanying consequences, complications such as infection and bleeding, the laboratory procedure and the risk of swapping-and laboratory-mistakes.		
◦ The chances of success, pregnancy after SET and DET, the chances of congenital diseases/malformations, ectopic pregnancy, multiple pregnancy and spontaneous abortion and possible unknown long-term risks.		
◦ The emotional aspects of the treatment, how patients can communicate that they are in need of extra emotional support and the existence of the patient association for infertility to share experiences and gain information from the patient's perspective.		
◦ How hospital care during treatment is organized and which caretakers are involved in the treatment procedure, how patients should contact the fertility team in case of problems or complaints and how the fertility team can be contacted, especially outside daytime-hours.		

WHO, World Health Organization; NVOG, derived from a guideline of the Dutch Society for Obstetrics and Gynaecology; NICE, (partially) derived from the Clinical Guideline 11 regarding infertility of NICE; Panel, item that was additionally suggested by the expert panel; SET, single embryo transfer; DET, double embryo transfer; VCM, volume x concentration x motility.

Discussion

When we consider our final set, it appears that 40% of the structure items and 50% of patient-information items were selected, which suggests that professionals acknowledge the importance of an accurate quality framework and complete patient information as conditions for 'best possible care'. Furthermore, the panellists suggested 36 new items before step 3, but only one of them was included in the final set (Table II). Experts apparently consider the current guideline contents as possibly incomplete, but still sufficiently satisfactory regarding clinical relevance.

The fact that the quality indicators in our set refer mainly to the process of care, a few to structure, and none to outcome reflects the ratios of these types of recommendations in guidelines. The lack of evidence for the best outcome measures in clinical care explains the absence of outcome-related recommendations in guidelines that are consensus and evidence based. Nevertheless, outcome indicators (e.g. 'live birth rate per treatment cycle') are the mean criterion used for judging professionals and hospitals in daily practice.[13] Outcome measures, on the one hand, have the disadvantage of being sensitive to case mix and the methods of data collection.[26,27] In addition, they may affect our judgment of the quality of care provided because good outcomes do not necessarily mean that care was delivered well. Process indicators, on the other hand, assess the process of care and reveal exactly where changes can be made, while structure indicators provide us with the framework to do so. Therefore, process and structure indicators are more valuable in quality improvement programmes because, unlike outcome measures, they offer a concrete starting point for improvement. This is especially true when process measures have been proven to relate to outcome measures.

In reproductive health care, choosing the best outcome measures to define the standards of success is currently being discussed internationally.[13,15,28,29] A comparison of this ongoing discussion and activity about outcome measures with the almost complete lack of initiatives for developing and measuring process and structure indicators, highlights a striking discrepancy. Apparently, little effort has been made to bridge the gap between the evidence we make available in our guidelines and the supposedly evidence-based care we deliver in practice.

The literature search revealed one similar initiative of indicator development for clinical fertility care, the NICE fertility guideline.[25] NICE used a systematic approach similar to ours to develop their auditable standards (http://www.nice.org.uk). However, a comparison of the contents of our final indicator set with NICE's seven key priorities shows that only one indicator is common to both sets: 'no more than two embryos should be transferred per embryotransfer' (Table II). Aside from similarities in guideline contents and health systems in both countries, different stakeholders and slightly different selection criteria have probably influenced the composition of both final sets. Especially the NICE criterion 'make efficient use of National Health Service resources', already curtails potential indicators because the lack of suitable resources would render indicators ineligible. In addition, NICE explicitly tries to include between 5 and 10 key priorities in their guidelines. This upper limit is in line with most quality

measure projects currently being realized.[30] Regarding fertility care, having many indicators would, in our opinion, better suit such a wide guideline programme that encompasses several different domains. However, before cutting back our set of 39 indicators, we recommend to perform a practice test.

There are some methodology limitations that could be considered with respect to the range of our final set. Comparable combined RAND-modified Delphi techniques have proven effective in several indicator-development initiatives.[20,31,32] However, the limitations and validity of such procedures have been questioned regularly.[33] In particular, the accuracy and validity is suggested to depend on several expert characteristics.[34] Questions may arise about whether the group of experts is heterogeneous enough and whether the individual expert contributions are of equal value. However, the group of experts in a small country like the Netherlands (which has only 13 licensed IVF clinics including eight university centres) is limited. We therefore think that the level of professional participation in our expert panels does reliably reflect the country's overall level of expertise. Moreover, a consensus-based-panel approach has the advantage of directly involving the target group of professionals in the procedure. By obtaining their consent during the selection process, we founded a substantial supporting base for our set of quality indicators, and consequently, for any national measurement programme or guideline-implementation strategy that may follow.

The systematic approach and panel composition underline how face and content validity are accounted for in this study. However, content validity alone is not enough to entitle indicators as valid. As already stated, improvement programmes in health care should preferably be tailor-made and therefore based on adequate measurements of performance. Although we have already considered selected criteria for operationalization (e.g. measurability and potential for improvement) during the selection rounds, the suggested quality indicators need to be submitted to a practice test. Then certain clinimetric characteristics, such as measurability, feasibility, reliability, and improvement potential, can be properly assessed. Furthermore, efforts should be made to try to relate indicator scores to outcome measures in order to assess the clinical importance of the separate quality indicators. Ultimately, this practice test will show whether these indicators will hold up in the future. Independently of any future adjustments, the complete indicator set presented in this paper complements the existing guidelines and can still serve as a benchmarking instrument for educational or quality-improvement activities in fertility centres.

This study took international literature into account, but still took place within the framework of the Dutch fertility guidelines. This implies that some of the selected indicators are particularly valuable in the Dutch situation and less suitable internationally. For example, the recommendation 'no more than two embryos should be transferred per embryotransfer' would probably not withstand the selection if the panellists were from the United States. Still, the topic of this indicator. e.g. 'the maximum number of embryos transferred' could be maintained for international application when adapted to specific national standards.

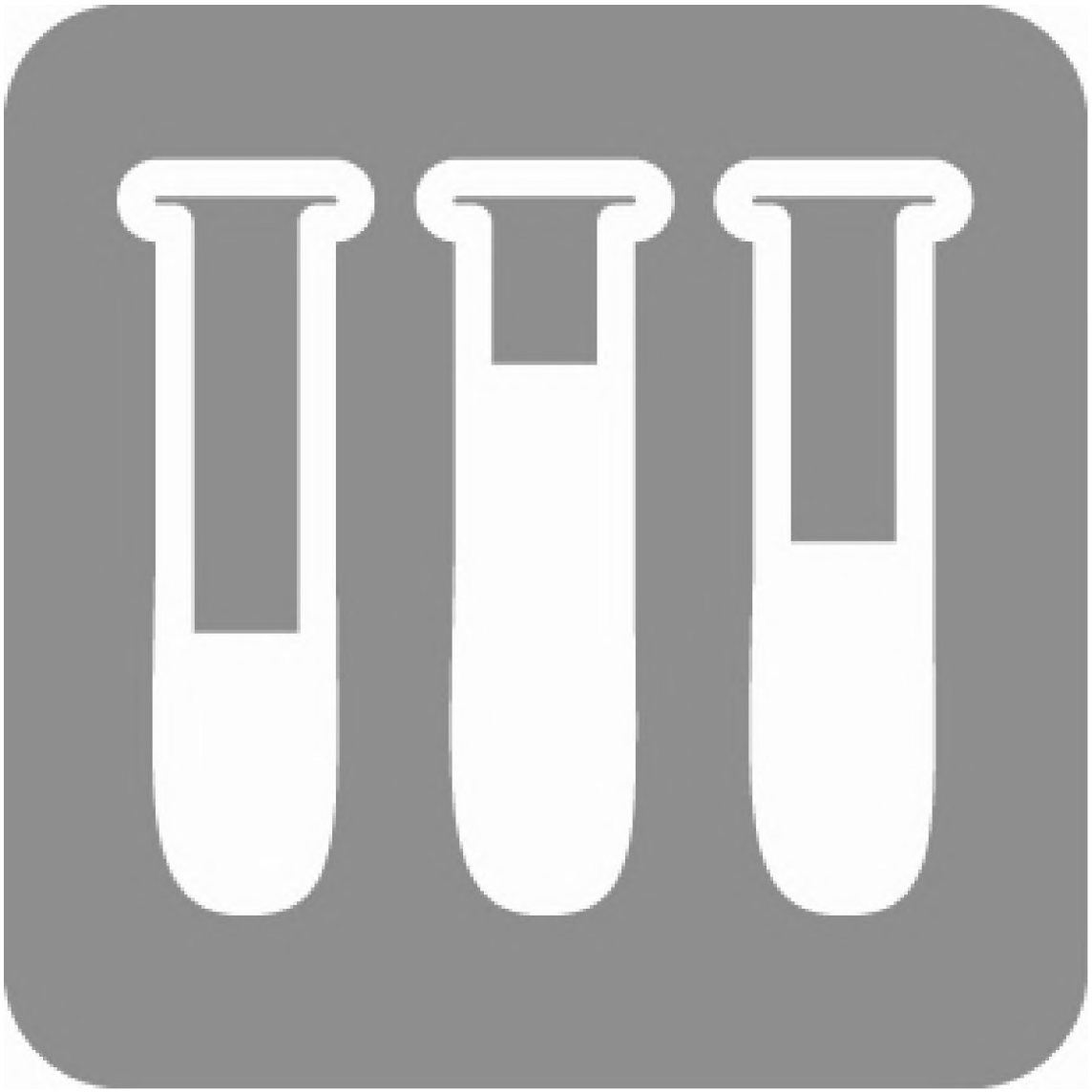
We strongly encourage the development of international fertility guidelines based on both international evidence and consensus, in initiatives like the one ESHRE is currently undertaking (www.eshre.com). However, we suggest that all such initiatives be accompanied by the simultaneous development of quality indicators by an international panel. Such quality indicators should be tested in practice first, and should preferably match the health-care systems and ethics of the various countries involved. Suitable data resources should be appointed or adapted to match the demands of proper indicator measurement. The consensus method and the domains of fertility care described in this paper could serve as a framework for further work in this field.

In conclusion, we have described a systematic procedure of developing a complete set of process and structure indicators based on an entire national fertility guideline programme. These 39 quality indicators can be used to monitor fertility care. However, their potential to do so accurately must first be proven in a proper practice test.

References

1. Vayena E, Rowe PJ, Griffin P. Current practices and controversies in assisted reproduction. Report of a WHO meeting. Geneva: World Health Organisation, 2002.
2. Nachtigall RD. International disparities in access to infertility services. Fertil Steril 2006;85:871-5.
3. Greil AL. Infertility and psychological distress: a critical review of the literature. Soc Sci Med 1997;45:1679-704.
4. Cousineau TM, Domar AD. Psychological impact of infertility. Best Pract Res Clin Obstet Gynaecol 2007;21:293-308.
5. Kallen B, Finnstrom O, Nygren KG, Otterblad OP, Wennerholm UB. In vitro fertilisation in Sweden: obstetric characteristics, maternal morbidity and mortality. BJOG 2005;112:1529-35.
6. Ombelet W, Martens G, De SP, Gerris J, Bosmans E, Ruyssinck G, Defoort P, Molenberghs G, Gyselaers W. Perinatal outcome of 12,021 singleton and 3108 twin births after non-IVF-assisted reproduction: a cohort study. Hum Reprod 2006;21:1025-32.
7. Schieve LA, Cohen B, Nannini A, Ferre C, Reynolds MA, Zhang Z, Jeng G, Macaluso M, Wright VC. A Population-Based Study of Maternal and Perinatal Outcomes Associated with Assisted Reproductive Technology in Massachusetts. Matern Child Health J 2007; in press.
8. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. BMJ 1999;318:527-30.
9. Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD, Thomson MA. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. The Cochrane Effective Practice and Organization of Care Review Group. BMJ 1998;317:465-8.
10. Grol R. Personal paper. Beliefs and evidence in changing clinical practice. BMJ 1997;315:418-21.
11. Grol R, Baker R, Moss F. Quality improvement research: understanding the science of change in health care. Qual Saf Health Care 2002;11:110-1.
12. Campbell SM, Braspenning J, Hutchinson A, Marshall MN. Research methods used in developing and applying quality indicators in primary care. BMJ 2003;326:816-9.
13. Min JK, Breheny SA, MacLachlan V, Healy DL. What is the most relevant standard of success in assisted reproduction? The singleton, term gestation, live birth rate per cycle initiated: the BESST endpoint for assisted reproduction. Hum Reprod 2004;19:3-7.
14. Nelen WL, Hermens RP, Mourad SM, Haagen EC, Grol RP, Kremer JA. Monitoring reproductive health in Europe: what are the best indicators of reproductive health? A need for evidence-based quality indicators of reproductive health care. Hum Reprod 2007;2:916-8.
15. Jahn A, Bloemenkamp KW, Hannaford P, Olsen J, da Silva MO, Temmerman M. Monitoring reproductive health in Europe--what are the best indicators of reproductive health? Hum Reprod 2006;21:2199-200.
16. Dalkey N, Brown B, Cochran S. The Delphi Method III: Use of Self Ratings to improve Group Estimates. 1969.
17. Fitch K, Bernstein S, Aguilar M. The RAND/UCLA Appropriateness Method User's Manual. Santa Monica, CA: RAND, 2001.
18. Campbell SM, Shield T, Rogers A, Gask L. How do stakeholder groups vary in a Delphi technique about primary mental health care and what factors influence their ratings? Qual Saf Health Care 2004;13:428-34.
19. WHO. The Constitution of the World Health Organization. 1946.
20. Campbell SM, Cantrill JA, Roberts D. Prescribing indicators for UK general practice: Delphi consultation study. BMJ 2000;321:425-8.
21. WHO. Monitoring reproductive health: selecting a short list of national and global indicators. 1997.
22. Ford J, Nassar N, Sullivan EA, Chambers G, Lancaster P. Reproductive health indicators Australia. 2003.
23. ICSI (2004) the Institute for Clinical System Improvement:The diagnosis and management of infertility (http://www.icsi.org/infertility/diagnosis_and_management_of_infertility_2301.html).
24. Temmerman M, Foster LB, Hannaford P, Cattaneo A, Olsen J, Bloemenkamp KW, Jahn A, da Silva MO. Reproductive health indicators in the European Union: The REPROSTAT project. Eur J Obstet Gynecol Reprod Biol 2006;126:3-10.
25. NICE. Fertility: assessment and treatment for people with fertility problems (Clinical guideline 11)

(http://www.nice.org.uk/pdf/CG011niceguideline.pdf). 2004.
26. Eddy DM. Performance measurement: problems and solutions. Health Aff (Millwood) 1998;17:7-25.
27. Mant J. Process versus outcome indicators in the assessment of quality of health care. Int J Qual Health Care 2001;13:475-80.
28. Schieve LA, Reynolds MA. What is the most relevant standard of success in assisted reproduction?: challenges in measuring and reporting success rates for assisted reproductive technology treatments: what is optimal? Hum Reprod 2004;19:778-82.
29. Dickey RP, Sartor BM, Pyrzak R. What is the most relevant standard of success in assisted reproduction?: no single outcome measure is satisfactory when evaluating success in assisted reproduction; both twin births and singleton births should be counted as successes. Hum Reprod 2004;19:783-7.
30. McGory ML, Shekelle PG, Rubenstein LZ, Fink A, Ko CY. Developing quality indicators for elderly patients undergoing abdominal operations. J Am Coll Surg 2005;201:870-83.
31. Schouten JA, Hulscher ME, Wollersheim H, Braspenning J, Kullberg BJ, van der Meer JW, Grol RP. Quality of antibiotic use for lower respiratory tract infections at hospitals: (how) can we measure it? Clin Infect Dis 2005;41:450-60.
32. Gagliardi AR, Fung MF, Langer B, Stern H, Brown AD. Development of ovarian cancer surgery quality indicators using a modified Delphi approach. Gynecol Oncol 2005;97:446-56.
33. Jones J, Hunter D. Consensus methods for medical and health services research. BMJ 1995;311:376-80.
34. Campbell SM, Hann M, Roland MO, Quayle JA, Shekelle PG. The effect of panel membership and feedback on ratings in a two-round Delphi survey: results of a randomized controlled trial. Med Care 1999;37:964-8.



Chapter 4

Variation in fertility care measured by guideline-based quality indicators

S.M. Mourad
W.L.D.M. Nelen
R.P.M.G. Hermens
L.F. Bancsi
D.D.M. Braat
G.A. Zielhuis
R.P.T.M. Grol
J.A.M. Kremer

Human Reproduction 2008; 23(11): 2493-500.

Abstract

Background: About 30–40% of patients do not receive care based on available scientific evidence. For infertility, this may imply unnecessary and expensive diagnostic tests and treatments. It is therefore important to identify gaps in performance by monitoring current fertility care. A set of 39 guideline-based quality indicators was previously developed for this purpose. This study aimed to assess several quality criteria of the indicator-set and to use the set to assess current fertility care.

Methods: A historic cohort study was performed in 16 Dutch fertility clinics; 2698 couples were invited to participate. Indicator data were gathered by medical record extraction, and patient and professional questionnaires. Quality criteria for each indicator (measurability, reliability, applicability, improvement potential, discriminatory capacity, complexity and case-mix stability) were assessed. Current practice was measured as adherence to the separate indicators.

Results: One thousand four-hundred and ninety-nine (56%) couples participated. All indicators were measurable, but the results for the other quality criteria varied. In total, 14 of the 39 indicators scored <50% adherence. Variation in performance between the clinics was up to 100%. The highest median adherence (86%) is found within the guideline 'indications for IVF-treatment'. The lowest median adherence is found within the guideline 'initial assessment of fertility' (43%), followed closely by the guideline 'anovulation' (44%).

Conclusions: This study shows the quality of the developed indicator-set for monitoring clinical fertility care. A first assessment in the Netherlands reveals large variation between clinics and ample room for improvement of care.

Introduction

According to medical literature, ~30–40% of patients do not receive care based on available scientific evidence.[1] On top of this, an estimated 20–25% of provided health care is unnecessary.[2-4] In the case of infertility, this could mean the use of unnecessary and expensive diagnostic tests and assisted reproduction technologies (ART) or the realization of complicated high-order pregnancies. Subsequently, this may have substantial physical and psychological consequences for the patients involved.[5-9] Patients should receive the best achievable care with a minimum of complications and at minimal costs. To improve both outcome and process of health care, clinical practice guidelines have been developed summarizing the best available evidence.[10] However, such guidelines do not implement themselves.[11] Efforts to monitor actual health care and to identify gaps in performance show where improvement is needed. Guideline-based quality indicators are useful tools for monitoring health care.[12] Consequently, they can be used to assess changes in clinical practice as well, which makes them indispensable instruments within clinical improvement programmes.

We reported previously on the development of a guideline-based set of 39 process and structure indicators for the entire spectrum of clinical fertility care by means of a systematic six-step RAND-modified Delphi method.[13] The selected indicators focus on process and structure of fertility care rather than outcome. Although current international debate focuses mainly on outcome measures[14,15], process indicators are valuable instruments within performance assessment and improvement programmes because they reveal exactly where changes in care might be necessary.[13,16-21] This does not mean, however, that clinical outcomes are unimportant, but to change outcomes, it is necessary to first initiate changes in the process and structure of care. The primary aim of this study was to test the indicator-set for several quality criteria.[22] Such a practice test is necessary to demonstrate its value as an instrument for monitoring and improvement of clinical performance.[23-27] The second aim, conditional on sufficient quality of the instrument, was to assess the variation in current fertility care in a large sample of clinics.

Materials and Methods

Study design and study population

We conducted a retrospective cohort study in 16 Dutch clinics using medical record and questionnaire data. The study was approved for all clinics by 'the Regional Review Board for Human Research (CMO) Arnhem-Nijmegen (CMO no. 2004/193)'. To assess current clinical fertility care in the Netherlands, we aimed at including a broad patient cohort from the participating clinics. The clinics' characteristics varied to ensure that delivered care was representative for Dutch standards. There was one academic and one tertiary care clinic and seven clinics offered secondary care (IVF/ICSI-treatment). These clinics were all also teaching clinics and of large or intermediate size; the other seven clinics were smaller, non-teaching facilities. In total, 15 clinics were national health, and one of the smaller secondary care clinics was a private clinic.

To include a representative patient group, potential participating couples were randomly selected by means of each clinics' financial DBC (Diagnosis/Treatment Combination code) registration database. In this national registration, patients undergoing diagnostics or treatment for infertility are identified with a specific

Fertility-code (F-code). Couples were apt for inclusion if they had an active F-code anytime between 1 April and 30 June 2005. The assessment focused on the care they received in the period 1 January–30 June 2005. In each clinic, a random sample of eligible couples (50–500 depending on the size of the clinic) was invited to participate in the study. They were sent an informed consent form and a questionnaire: the former included consent to use their medical data from the clinic. Couples who had insufficient knowledge of the Dutch language or did not visit the fertility clinic in the study period were excluded. In total, 2698 couples were invited.

The clinical quality indicators

We used a set of 39 clinical quality indicators for fertility care, which was developed using a rigorous and systematic six-step approach. The set was based on literature, existing international guidelines and indicators and 10 Dutch (evidence and consensus based) fertility guidelines.[13] The latter includes the Dutch Embryo Act and nine fertility guidelines of the Dutch Society of Obstetrics and Gynecology (NVOG): initial assessment of fertility, anovulation and child wish, male infertility, tubal pathology, endometriosis, premature ovarian failure (POF), intrauterine insemination (IUI), indications for IVF treatment and ovarian hyperstimulation syndrome (OHSS).

The indicator-set encompasses 37 process indicators and 2 structural indicators, and comprises the entire spectrum of fertility care: the initial assessment of fertility, diagnostics, treatments and possible complications. Indicators were operationalized by the construction of a numerator and denominator, each consisting of several variables. The numerator is formed by the proportion of patients in which there is adherence to a recommendation and the denominator by the proportion of patients in which the recommendation is applicable. The indicators within the guideline 'initial assessment of fertility' are shown in Table I. The remaining indicators, ranked per guideline, are shown in Supplementary Table Ia–e.

Data collection to assess quality criteria and current practice

Data for indicators regarding indications for ART (n=7), diagnostic procedures (n=6) and treatment procedures (n=20) were extracted from the individual medical records by three trained investigators. Data for indicators regarding information provision to patients (n=4) were collected by means of a patient questionnaire and data for indicators regarding the structure of care (n=2) by a short professional questionnaire. The patient questionnaire was piloted in a group of 30 infertile couples and this led to minor adjustments in formulation of the questions before it was used in the total study population.

Definitions and analysis

The application of a systematic and rigorous consensus method for indicator development results in high face and content validity of this indicator-set. We measured the following quality criteria of all indicators in the set, in order to demonstrate its value as an instrument for assessment of fertility care.[12,23-25,27-30] Interested readers are referred to Supplementary Table II in the web-based version of this article which contains a list of suggested additional reading on the use and validity of (process) indicators in clinical quality and performance assessment.

(i) **Measurability:** an indicator is considered measurable if data to fill the numerator and denominator of the indicator can be made available through data collection (e.g. medical records, complication or treatment databases or a survey).[23]

(ii) **Reliability:** if the measurement of indicator data by two different data-collectors is reliably comparable, there is high reliability.[23] Two independent reviewers abstracted a random sample of 10% (n=32) of medical records from two participating clinics. The extent of agreement between these data reviewers on the level of process indicator scores, corrected for chance, was calculated using Cohen's kappa coefficient of inter-rater reliability. Coefficients of 0.4–0.6 represent moderate inter-rater reliability, coefficients of 0.61 or higher are considered very good.[31]

(iii) **Applicability:** for accurate performance assessment, an indicator is preferably applicable to a substantial proportion of reviewed patients (>10 cases); this is referred to as applicability.[12]

(iv) **Improvement potential**: when an indicator is used to detect changes in clinical performance, it is a prerequisite that improvement is possible at all; if overall performance for a certain indicator is already very high, the indicator has no improvement potential. We defined improvement potential as an overall performance score of <90%.[27]

(v) **Discriminatory capacity:** the discriminatory capacity indicates whether an indicator is able to discriminate practice performance between different hospitals. High discriminatory capacity is therefore present when the range in scores between the lowest- and highest-scoring clinics is >20%.[12]

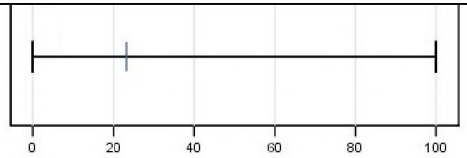
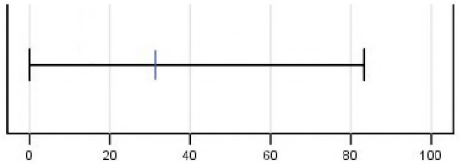
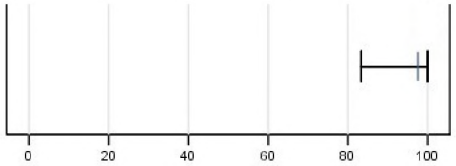
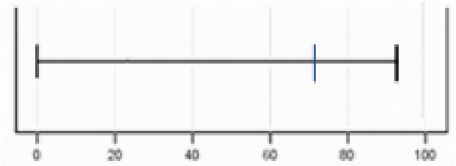
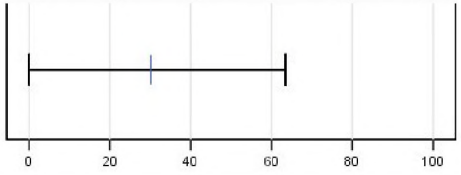
(vi) **Complexity:** the concept of complexity was used for the number of variables needed to fill the numerator and denominator of an indicator. Complexity is thus a measure that reflects the amount of investments needed to assess an indicator; the higher the complexity the more effort it takes to measure an indicator. The maximum value for complexity was beforehand set at five different variables; in the case of complexity scores of more than five variables, subdivision of the indicator was considered. Variables relating to the inclusion of patients were not counted because those variables can vary per clinic and depend on design of the medical records. For the indicators measured by medical record extraction, two researchers (W.L.D.M.N. and S.M.M.) counted independently the number of variables to fill the numerator and denominator of each indicator. For the indicators measured by means of a professional or patient questionnaire, we counted the number of answers needed to determine the indicator. If there was disagreement between the two researchers' findings, discussion took place to reach an agreement.

(vii) **Case-mix stability:** case-mix stability is an important asset that enables the application of an indicator for monitoring within a specific hospital over time or to compare hospitals of different sizes and settings.[23] Therefore, the relationship between certain patient-characteristics that may vary considerably between clinics ['woman's age', 'type of infertility (primary or secondary)' and 'duration of infertility'] and the indicator scores was analysed to decide whether correction for case mix is necessary.

Analysis

Collected data were entered in a database using the Statistical Package for the Social Sciences (SPSS 14.0 for Windows, SPSS Inc., Chicago, IL, USA). Descriptive analyses were performed for each indicator, and current practice was expressed as percentage of adherence to an indicator. The median adherence of each separate clinic was calculated, as well as the adherence of all clinics together.

Table I. Adherence to indicators for clinical fertility care: an assessment in 16 Dutch fertility clinics.

Initial assessment of fertility	Range and median adherence for 16 clinics (%)	Median adherence (43[a])	Range	*n* patients
The initial fertility assessment should result in both a diagnosis and a prognosis.		23	0-100	428
The initial fertility assessment should consist of three parts: semenanalysis, tubal patency assessment. and cycle-analysis.		31	0-83	179
Couple's history-taking should cover at least: 1. age of both partners, 2. duration of infertility, 3. type of couple's infertility (primary or secondary)		98	83-100	361
Woman's physical examination should include assessment of the Body Mass Index.		72	0-95	333
Life-style advice concerning bodyweight, smoking, and the alcohol and drug use should be part of the counselling regarding pregnancy-probabilities.		30	0-63	295

[a] median adherence of all indicators within one guideline

Table II. Baseline characteristics of the participating couples (n=1499)

Characteristics		% of couples
Mean age in years (SD)		
Female	32.8 years (4.1)	
Male	35.1 years (5.0)	
Ethnic background[a]		
Dutch		98.1
Non-Dutch		1.9
Gross monthly family income (euros)[b]		
<1100		1.6
1100–1760		4.6
1760–2750		22.2
>2750		71.6
Education level per couple[c]		
Low		6.4
Intermediate		40.7
High		52.9
Type of infertility[d]		
Primary		72.8
Secondary		27.2
Mean duration of infertility in months (SD)[e]		38.4 (22.5)
BMI of the woman		
Low (<18)		14.0
Normal (18–25)		53.9
High (>25)		32.0

[a] Ethnic background of the couples was determined by the origin of both partners: Dutch, one or both partners are of Dutch origin; non-Dutch, both partners are not of Dutch origin.

[b] Gross monthly family income was categorized according to social security standards and modal income in euros: ,1100 less than Dutch modal income; <1100–1760 Dutch modal income; 1760–2750 up to 1.5 x Dutch modal income; .2750 more than Dutch modal income.

[c] Education level of the couples was determined by the highest education level of both partners: low, primary or lower vocational education; intermediate, secondary or intermediate vocational education; high, higher professional education or university.

[d] Type of fertility was determined for the couple.

[e] Duration of infertility was defined as the period between the start of regular unprotected sexual intercourse and 1 January 2005, the beginning of the study period.

Table III. Quality criteria of guideline-based indicators for clinical fertility care: an assessment in 16 Dutch fertility clinics.

Initial assessment of fertility	Applicability ($n_{patients}$)	Improvement potential	Discriminatory capacity (%)	Complexity ($n_{variables}$)	Case mix stability	Data source
The initial fertility assessment should result in both a diagnosis and a prognosis.	428	Y	100	2	Y	Medical record
The initial fertility assessment should consist of three parts: semen-analysis, tubal patency assessment. and cycle-analysis.	179	Y	83	3	N[a]	Medical record
Couple's history-taking should cover at least: 1. age of both partners, 2. duration of infertility, 3. type of couple's infertility (primary or secondary)	361	N	17	5	Y	Medical record
Woman's physical examination should include assessment of the Body Mass Index.	333	Y	95	2	Y	Medical record
Life-style advice concerning bodyweight, smoking, and the alcohol and drug use should be part of the counselling regarding pregnancy-probabilities.	295	Y	63	4	N[a]	Patient questionnaire

Quality indicators of the guideline "Initial assessment of fertility"; N=no, Y=yes, NA= not applicable
[a]case mix correction is necessary for 'duration of infertility'.

Results

Response

The mean participation rate was 56% (1499 of 2698 invited couples), varying between 47% and 72% per clinic. Demographic characteristics of the participating couples are shown in Table II. Mean female age was 33 years and male age 35 years. Forty-seven couples (3.2%) were of non-Dutch origin. Of the participating couples, 73% suffered from primary infertility. The median duration of infertility was 38 months. A total of 53% had a high education level and 72% had a more than modal income (gross > € 2750/month).

Assessment of quality criteria of the indicator-set

Table III shows the quality criteria of the indicators, within the guideline 'initial assessment of fertility'. Supplementary Tables IIIa–i show these criteria for the remaining guidelines. All indicators were measurable and made available through either medical record extraction, or a patient or professional questionnaire. The reliability is high for questionnaire data, because it is first-hand information. For medical record data, reliability was substantial, reflected in an average Cohen's kappa coefficients of 0.86 (range 0.48–1.0). In total, 18 indicators were inapplicable (<10 patients). For these indicators, the other quality criteria could not be assessed. A total of eight indicators had adherence scores of 90% or more, and therefore possess too little room for improvement. Regarding discriminatory capacity, a total of 11 indicators scored <20%. This understandably includes the eight indicators that had no room for improvement. None of the indicators required more than one data source. Complexity values ranged mostly from 1 to 5. An exception was the indicator concerning 'the items that should be discussed during an intake prior to IVF/ICSI-treatment'. This indicator encompassed many topics and concerned 18 variables and was therefore subdivided into five different parts for accurate analysis. Assessment of case-mix stability revealed that the following three indicators are in need of correction for 'duration of infertility': (i) 'the initial fertility assessment should consist of three parts: semen-analysis, tubal patency assessment and cycle analysis', (ii) 'life-style advice concerning bodyweight, smoking and the alcohol and drug use should be part of the counseling regarding pregnancy-probabilities' and (iii) 'in the case of IUI in the stimulated cycle, ovarian response should be monitored by transvaginal ultrasound'. The indicator 'in the case of unexplained infertility in a woman, 36 years, there is an indication for IVF after 3 years of infertility' needs correction for 'duration of infertility', 'woman's age' and 'type of infertility' when the analysis takes place. All other indicators showed equal distribution of these patient-characteristics. The overall results of this quality criteria assessment are summarized in Table IV.

Table IV. Overall scores of the indicators (n=39) for each quality criterion.

Quality criteria	Number of indicators that met the criterion
Measurability	39
Reliability	39
Applicability	21
Improvement potential	31
Discriminatory capacity	28
Complexity	38
Case-mix stability	35

Current practice: adherence to indicators

Table I shows the median clinical performance and the range of performance of the clinics for each indicator within the guideline 'initial assessment of fertility'. Supplementary Table Ia–e contains the results for the remaining indicators. The indicators regarding endometriosis, OHSS, tubal pathology and POF [13] were applicable to too few patients (n<10) in this sample and are therefore not shown in Table I. From Table I and Supplementary Table Ia–e, it can be concluded that the median adherence of 14 indicators was <50%. Of those, six indicators even scored <25%. The median adherence to the separate indicators varied from 0% (e.g. 'in the case of abnormal semen-analysis, a complete andrological history taking, physical examination and one extra semen-analyses should be performed') up to 100% (e.g. 'Each department performing IUI should evaluate their results annually').

The variation between clinics differed widely and is reflected in the range of adherence within one indicator. We found, for example, on the one hand exactly equal adherence between all clinics (e.g. all clinics scored 100%) for the indicator 'routine use of hCG for luteal support after IVF is not recommended' and on the other hand a 100% difference between the clinics (e.g. two clinics scoring 0% and 100%, respectively) for the indicator 'initial fertility assessment should result in a prognosis and a diagnosis'.

Finally, the results show that the highest median adherence (86%) is found within the guideline 'indications for IVF-treatment'. The lowest mean adherence is found within the guideline 'initial assessment of fertility' (43%), followed closely by the guideline 'anovulation' (44%).

Discussion

This study demonstrates several quality criteria of a set of 39 guideline-based process and structure indicators for comprehensive clinical fertility care. Moreover, a first assessment of current fertility care in the Netherlands discloses a large variation in the performance between clinics.

During the original indicator-selection procedure, the expert panels were asked to take five criteria into account when selecting key recommendations from the guidelines.[13] In this practice test, we assessed additional quality criteria of the proposed indicators. We can conclude that most of them meet these criteria, with several remarkable results. Improvement potential and applicability were not always high. Apparently, a considerable discrepancy exists between the experts' estimate of these criteria on the one hand and currently provided care on the other hand. This underlines that even a rigorous indicator-development procedure should be complemented with a practice

test. On the basis of the results of our practice test, particularly applicability and improvement potential (Tables III and IV and Supplementary Table IIIa–i), we can distinguish three subsets within our original set of 39 quality indicators. Indicators within the 'first subset' score highly for applicability and improvement potential and can therefore be used for quality monitoring purposes and as a baseline measurement for improvement programmes in the involved clinics. A 'second subset' contains indicators with low improvement potential, which makes them useless for improvement programmes, but with high applicability, which makes them still suitable for monitoring purposes, namely to ensure that adherence continues to be high in the future. Within the 'third subset', we find indicators with low applicability. They concern relatively rare conditions or complications (e.g. POF and severe OHSS), so inclusion numbers within our random patient sample were too small for adequate indicator assessment. This does not mean, however, that this subset needs to be discarded immediately for any future use; but the inclusion of enough cases for such indicators requires a more specific patient sample, e.g. drawn from separate complication registrations or large (national) databases.
The question arises whether this set of indicators is also suitable for countries other than the Netherlands. Only 2 out of the 39 indicators (e.g. laboratory accreditation by a Dutch committee and the maximum number of two embryos transferred per cycle) may be more or less specific for Dutch practice, but both can be easily adapted for international use by judging not the content but the topic of the indicator. Generalizability is thus not a barrier for international adaptation of the presented indicator-set.

Adherence to the guideline recommendations
Our study shows that adherence to guideline-based indicators ranged widely per clinic, per guideline and per individual indicator. Adherence to the indicators is relatively poor; 14 indicators score below 50% adherence. This means there is ample room for improvement of clinical fertility care. Professionals should be urged to improve the implementation of those guideline-recommendations that showed poor adherence in their clinic. A 100% adherence score may not always be a viable goal, because there can in specific cases be good reasons to divert from guideline-recommendations. However, professionals should at least aim to rival the best scoring clinic, which is a realistic benchmark. The lowest adherence (less than 50%) is found within the guidelines 'initial assessment of fertility' and 'anovulation'; priority should therefore be given to implementation of these guidelines. It is furthermore important to realize that a low score on an indicator does not automatically mean that there is a problem in the quality of care; it is, however, a signal to further estimate the matter in order to try to understand underlying causes and processes.

There are some results that stand out in particular. Several of the critical indicators with low adherence relate to communication and patient information. Again, the broad variation in adherence to these indicators and the fact that some clinics do score highly, illustrate that these results cannot be simply disregarded or attributed to recall bias.[32,33] These findings are in line with previously conducted qualitative and quantitative research, identifying physicians' 'lack of self-efficacy regarding

communication' and 'low outcome expectancy' as main barriers for adherence to a fertility guideline.[34] However, this is plainly disturbing from the perspective of patient-centred care, as former research clearly discloses infertile patients' preferences for comprehensible and complete information.[35-38] Increased awareness among careproviders of these patient preferences and feedback on low indicator scores should lead to an improvement in information provision and transparency of care. Another low-scoring indicator is the one recommending unstimulated over stimulated IUI, in the case of unexplained infertility, to prevent multiple pregnancies (15% adherence, range between clinics 0–48%). This perfectly reflects the situation where consensus on a specific topic is lacking. None of the clinics score 100% adherence, meaning this recommendation is far from consistently followed. This might be because both patients and professionals are apt to change their treatment policy and are willing to take more risks, e.g. after previous unsuccessful cycles of unstimulated IUI. This probably also explains the poor adherence to the recommendation 'mono-ovulation should be the result of ovulation-induction'. However, it is the responsibility and moral duty of the professional to guard patient safety during any treatment, and 0% scores should therefore be taken very seriously. Fortunately, more reticence is seen in the high adherence (79%) of the recommendation 'IVF-treatment in women <36 years is only indicated after 3 years of infertility'. This reflects the professionals' commitment to prevent over-treatment and strive for cost-effectiveness and efficacy of this expensive and intensive type of infertility treatment, by respecting the possibility of less invasive approaches and treatment-independent pregnancies.[39,40]

This study also has some limitations. Indicator measurement turned out to be quite a laborious exercise, including widely distributed questionnaires and an extensive medical record search to complete indicator data. Poor availability of medical data and the lack of adequate data resources are common problems in performance assessment efforts[41,42], but this should never be the sole reason to discard rigorously selected quality indicators. On the contrary, care-providers and policy-makers should undertake efforts to overcome these barriers. The ongoing introduction of electronic patient records offers great opportunities for collaboration with clinical improvement programmes. It is therefore important to conduct further research to uncover which relatively small investments or adjustments to existing databases are necessary to facilitate monitoring of current care.

In conclusion, we demonstrated the quality of a developed set of guideline-based indicators for comprehensive clinical fertility care. A practice test in the Netherlands revealed large variation and ample room for improvement of fertility care. Such objective assessment of care can help professionals to identify and subsequently target the domains of care in need of improvement. The next step will be the translation of indicator results into appropriate implementation strategies that aim to bridge the gaps between the best available evidence we present in our guidelines and the care we deliver in daily practice.

Table Ia. Adherence to indicators for clinical fertility care: an assessment in 16 Dutch fertility clinics.

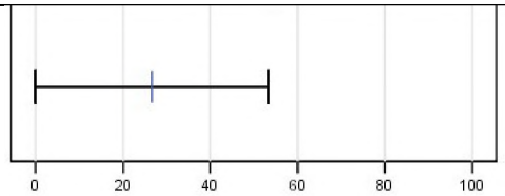
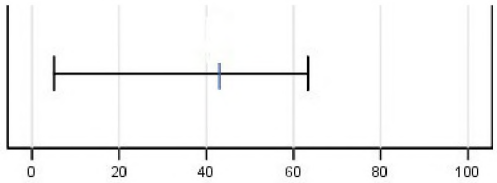
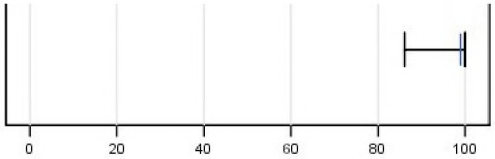
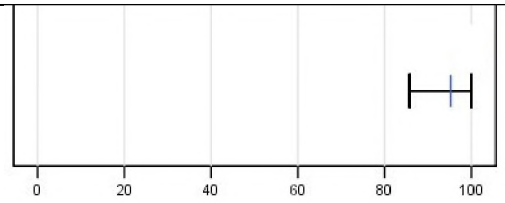
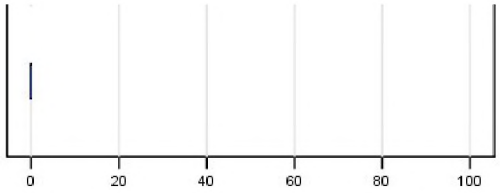
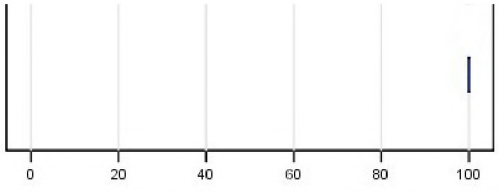
	Range and median adherence for 16 clinics (%)	Median adherence	Range	n patients
Anovulation and Childwish		**44** [a]		
Patients with overweight should, with regard to their fertility-treatment and overall health, be informed of the importance of weight reduction by means of life-style changes.		27	0-53	126
The goal of ovulation-induction should be mono-ovulation.		43	5-63	52
The ovarian response to hormonal stimulation should be performed by regular transvaginal ultrasound (frequency of 1-3 times / week).		86	86	30
Women with WHO group II anovulatory infertility should be given anti-oestrogen as first-choice medication for ovulation induction.	NA	NA	NA	2
Before starting ovulation induction treatment, the patient should be informed about: 1. the specific side-effects of medication 2. the need for regular intensive follow-up during treatment 3. the increased risks of multiple pregnancy 4. the ovarian hyper stimulation syndrome 5. a spontaneous abortion.		19	0-50	202
In case of 3 or more follicles >16mm and/or more than 5 follicles>12mm during an ovulation-induction treatment cycle, the patient has to be informed that coitus is prohibited or contraception has to be used.	NA	NA	NA	9

Table Ib. Adherence to indicators for clinical fertility care: an assessment in 16 Dutch fertility clinics.

	Range and median adherence for 16 clinics (%)	Median adherence	Range	n patients
Male factor infertility		**65** [a]		
In case of normospermia (WHO-criteria), semen-analysis should not be repeated, and no additional andrological tests should be performed.		95	86-100	78
In case of an abnormal semen-analysis (WHO-criteria), the physician should perform: 1. a complete andrological history-taking 2. a physical examination 3. at least one extra semen-analysis. *(In case of partial andrological history-taking)*[d]		0	0	59
In case of an idiopathic oligoasthenoterato-zoospermia, no hormones, vitamins or NSAID's should be prescribed to improve semen-quality.		100	0	179

[a] median adherence of all indicators within one guideline; NA= not applicable; WHO = World Health Organisation; NSAID= non steroidal anti-inflammatory drugs

Table Ic. Adherence to indicators for clinical fertility care: an assessment in 16 Dutch fertility clinics.

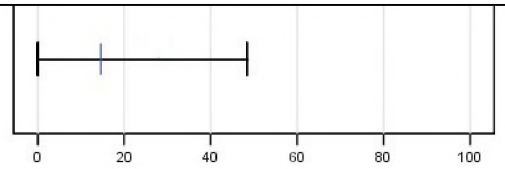
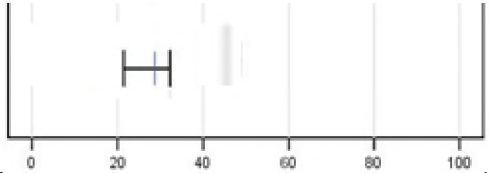
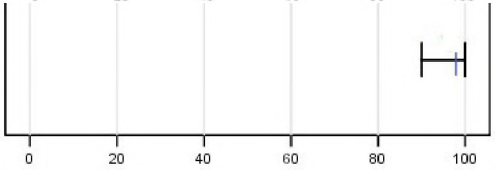
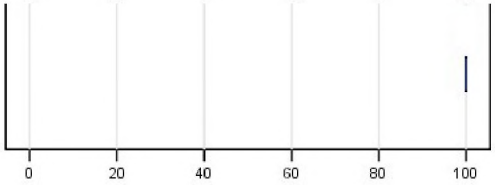
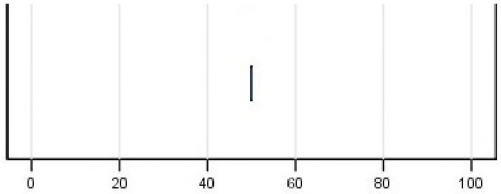
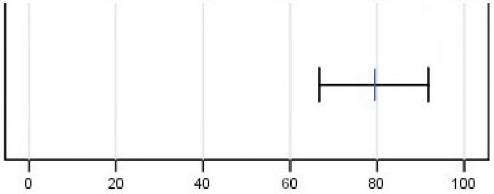
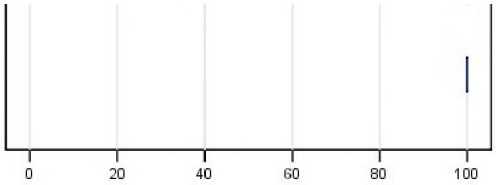
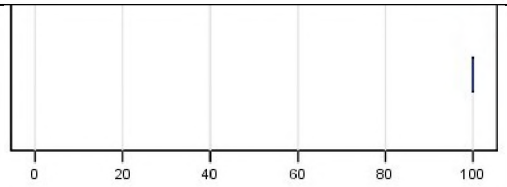
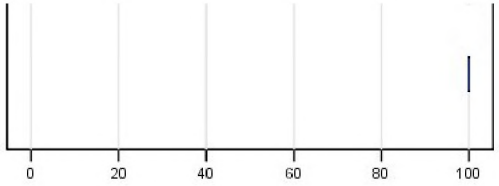
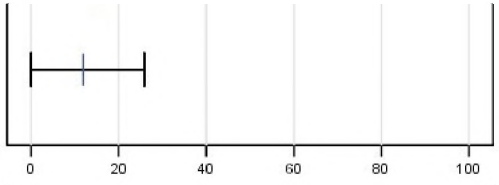
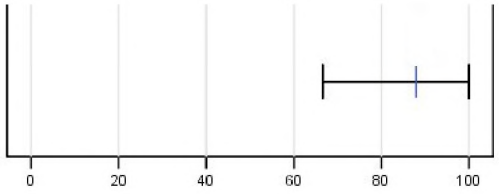
	Range and median adherence for 16 clinics (%)	Median adherence	Range	n patients
Intra uterine insemination (IUI)		60 [a]		
In case of unexplained infertility, stimulated IUI should not be offered, even though it is associated with higher pregnancy rates than unstimulated IUI, because it carries a risk of multiple pregnancy.		15	0-48	138
The diagnosis 'cervical factor' is an indication for IUI in the unstimulated cycle.		26	21-31	41
In case of IUI in the stimulated cycle, ovarian response should be monitored by transvaginal ultrasound.		98	90-100	173
Each department performing IUI should evaluate their results annually.		100	0	16
IUI should not be performed in case of more than 3 follicles > 16 mm or more than 5 follicles > 12mm. In both cases, the use of contraceptives should be advised as well.	NA	NA	NA	9

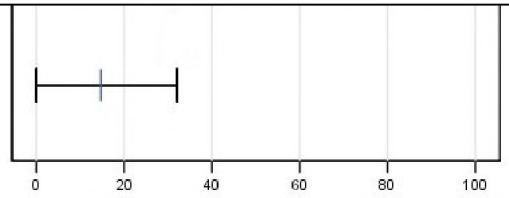
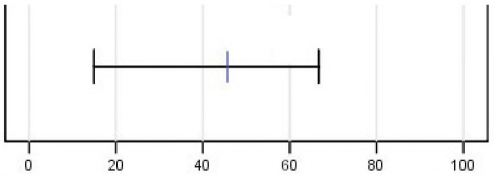
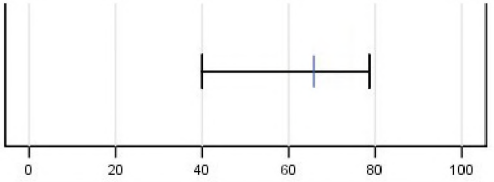
Table Id. Adherence to indicators for clinical fertility care: an assessment in 16 Dutch fertility clinics.

	Range and median adherence for 16 clinics (%)	Median adherence	Range	n patients
Indications for IVF/ICSI		**76** [a]		
In case of severe endometriosis and decreased tubal function (but no *bilateral* occlusion), there is an indication for IVF after 2 years of infertility.	NA	NA	NA	0
In case of male infertility with VCM<1x10*6/ml (before capacitation) there is a direct indication for ICSI-treatment		50	50	24
In case of unexplained infertility in a woman <36 years, there is an indication for IVF after 3 years of infertility.		79	67-92	43
Routine use of hCG for luteal support after IVF is not recommended.		100	100	16

[a] median adherence of all indicators within one guideline
NA= not applicable, VCM=volume*motility*concentration

Table Ie. Adherence to indicators for clinical fertility care: an assessment in 16 Dutch fertility clinics.

	Range and median adherence for 16 clinics (%)	Median adherence	Range	n patients
Embryo Act		**58** [a]		
IVF-Laboratories should be accredited by the CCKL (Dutch Coordinating Committee for maintenance of Quality Standards in Laboratory research).		100	100	2
Per embryo-transfer no more than 2 embryos should be transferred at the same time.		100	NA	265
During an intake prior to IVF/ICSI-treatment, the following should be discussed:				
1.The risks of hyperstimulation, poor response and accompanying consequences, complications such as infection and bleeding, the laboratory procedure and the risk of swapping- and laboratory-mistakes.		12	0-26	270
2.The chances of success, pregnancy after SET and DET.		88	67-100	300

	Range and median adherence for 16 clinics (%)	Median adherence	Range	n patients
Embryo Act		**58** [a]		
3.the chances of congenital diseases/malformations, ectopic pregnancy, multiple pregnancy and spontaneous abortion and possible unknown long-term risks.		15	0-32	256
4.The emotional aspects of the treatment, how patients can communicate that they are in need of extra emotional support and the existence of the patient association for infertility to share experiences and gain information from the patient's perspective.		46	15-67	280
5.How hospital care during treatment is organised and which caretakers are involved in the treatment procedure, how patients should contact the fertility team in case of problems or complaints and how the fertility team can be contacted, especially outside daytime-hours.		66	40-79	264

[a] median adherence of all indicators within one guideline; NA= not applicable

Table II. Suggested reading on the use and validity of (process) indicators in clinical quality and performance assessment.

AHCPR (1995) Using Clinical Practice Guidelines to evaluate Quality of Care. AHCPR Publication no. 95-0045. Rockville: US Department of Health and Human Services.

Arling G, Kane RL, Lewis T and Mueller C (2005) Future development of nursing home quality indicators. Gerontologist 45, 147-156.

Brook RH, McGlynn EA and Shekelle PG (2000) Defining and measuring quality of care: a perspective from US researchers. Int J Qual Health Care 12, 281-295.

Campbell SM, Braspenning J, Hutchinson A and Marshall MN (2003) Research methods used in developing and applying quality indicators in primary care. BMJ 326, 816-819.

Grol R, Wensing M and Eccles M (2005) Improving patient care. The implementation of change in clinical practice. Chapter 16: 'Measuring changes in patient care: development and use of indicators'. 1st edition, Elsevier Limited, London, UK.

Grol R, Baker R and Moss F (2002) Quality improvement research: understanding the science of change in health care. Qual Saf Health Care 11, 110-111.

Landis JR and Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33, 159-174.

Mant J (2001) Process versus outcome indicators in the assessment of quality of health care. Int J Qual Health Care 13, 475-480.

Palmer RH (1997) Process-based measures of quality: the need for detailed clinical data in large health care databases. Ann Intern Med 127, 733-738.

Rubin HR, Pronovost P and Diette GB (2001) From a process of care to a measure: the development and testing of a quality indicator. Int J Qual Health Care 13, 489-496.

Streiner D and Norman G (1994) Health measurement scales. 2nd. ed. Oxford University Press, Oxford, UK.

Wollersheim H, Hermens R, Hulscher M, Braspenning J, Ouwens M, Schouten J, Marres H, Dijkstra R and Grol R (2007) Clinical indicators: development and applications. Neth J Med 65, 15-22.

Table IIIa. Quality criteria of guideline-based indicators for clinical fertility care: an assessment in 16 Dutch fertility clinics.

Anovulation and Childwish	Applicability ($n_{patients}$)	Improvement potential	Discriminatory capacity (%)	Complexity ($n_{variables}$)	Case mix stability	Data source
Patients with overweight should, with regard to their fertility-treatment and overall health, be informed of the importance of weight reduction by means of life-style changes.	126	Y	53	3	Y	Patient questionnaire
The goal of ovulation-induction should be mono-ovulation.	52	Y	58	2	Y	Medical record
The ovarian response to hormonal stimulation should be performed by regular transvaginal ultrasound (frequency of 1-3 times / week).	30	Y	0	2	NA	Medical record
Women with WHO group II anovulatory infertility should be given anti-oestrogen as first-choice medication for ovulation induction.	2	NA	NA	1	NA	Medical record
Before starting ovulation induction treatment, the patient should be informed about: 1.the specific side-effects of medication 2.the need for regular intensive follow-up during treatment 3.the increased risks of multiple pregnancy 4.the ovarian hyper stimulation syndrome 5.a spontaneous abortion.	202	Y	50	5	Y	Patient questionnaire
In case of 3 or more follicles >16mm and/or more than 5 follicles>12mm during an ovulation-induction treatment cycle, the patient has to be informed that coitus is prohibited or contraception has to be used.	9	NA	NA	3	Y	Medical record

Quality indicators of the guideline 'Anovulation and Childwish'.

Table IIIb. Quality criteria of guideline-based indicators for clinical fertility care: an assessment in 16 Dutch fertility clinics.

Male factor infertility	Applicability ($n_{patients}$)	Improvement potential	Discriminatory capacity (%)	Complexity ($n_{variables}$)	Case mix stability	Data source
In case of normospermia (WHO-criteria), semen-analysis should not be repeated, and no additional andrological tests should be performed.	78	N	14	2	Y	Medical record
In case of an abnormal semen-analysis (WHO-criteria), the physician should perform: 1. a complete andrological history-taking 2. a physical examination 3. at least one extra semen-analysis.	59	Y	36	4	Y	Medical record
In case of an idiopathic oligoasthenoterato-zoospermia, no hormones, vitamins or NSAID's should be prescribed to improve semen-quality.	179	N	0	4	NA	Medical record

Quality indicators of the guideline 'Male factor infertility'.

[a] case mix correction is necessary for 'duration of infertility'; [b] case mix correction is necessary for 'duration of infertility', 'female age' and 'type of infertility for the couple'; N=no, Y=yes, NA= not applicable, NSAID=non steroidal anti-inflammatory drugs, WHO = World Health Organisation

Table IIIc. Quality criteria of guideline-based indicators for clinical fertility care: an assessment in 16 Dutch fertility clinics.

Intra uterine insemination (IUI)	Applicability ($n_{patients}$)	Improvement potential	Discriminatory capacity (%)	Complexity ($n_{variables}$)	Case mix stability	Data source
In case of unexplained infertility, stimulated IUI should not be offered, even though it is associated with higher pregnancy rates than unstimulated IUI, because it carries a risk of multiple pregnancy.	138	Y	48	1	NA	Medical record
The diagnosis 'cervical factor' is an indication for IUI in the unstimulated cycle.	41	Y	10	1	Y	Medical record
In case of IUI in the stimulated cycle, ovarian response should be monitored by transvaginal ultrasound.	173	N	10	1	N[a]	Medical record
Each department performing IUI should evaluate their results annually.	16[c]	N	0	1	NA	Professional questionnaire
IUI should not be performed in case of more than 3 follicles > 16 mm or more than 5 follicles > 12mm. In both cases, the use of contraceptives should be advised as well.	9	NA	NA	4	NA	Medical record

Quality indicators of the guideline 'Intra uterine insemination'.

Table IIId. Quality criteria of guideline-based indicators for clinical fertility care: an assessment in 16 Dutch fertility clinics.

Indications for IVF/ICSI	Applicability ($n_{patients}$)	Improvement potential	Discriminatory capacity (%)	Complexity ($n_{variables}$)	Case mix stability	Data source
In case of severe endometriosis and decreased tubal function (but no *bilateral* occlusion), there is an indication for IVF after 2 years of infertility.	0	NA	NA	4	NA	Medical record
In case of male infertility with VCM<1x10*6 /ml (before capacitation) there is a direct indication for ICSI-treatment	24	Y	0	3	Y	Medical record
In case of unexplained infertility in a woman <36 years, there is an indication for IVF after 3 years of infertility.	43	Y	25	5	N[b]	Medical record
Routine use of hCG for luteal support after IVF is not recommended.	16	N	0	1	NA	Medical record

Quality indicators of the guideline 'Indications for IVF/ICSI'.

[a] case mix correction is necessary for 'duration of infertility'; [b] case mix correction is necessary for 'duration of infertility', 'female age' and 'type of infertility for the couple'; [c] this indicator applies to clinical departments, not patients; N=no, Y=yes, NA= not applicable; VCM=volume*motility*concentration

Table IIIe. Quality criteria of guideline-based indicators for clinical fertility care: an assessment in 16 Dutch fertility clinics.

Embryo Act	Applicability ($n_{patients}$)	Improvement potential	Discriminatory capacity (%)	Complexity ($n_{variables}$)	Case mix stability	Data source
IVF-Laboratories should be accredited by the CCKL (Dutch Coordinating Committee for maintenance of Quality Standards in Laboratory research).	2[d]	N	0	1	NA	Professional questionnaire
Per embryo-transfer no more than 2 embryos should be transferred at the same time.	265	N	0	2	NA	Medical record
During an intake prior to IVF/ICSI-treatment, the following should be discussed:						
1. The risks of hyperstimulation, poor response and accompanying consequences, complications such as infection and bleeding, the laboratory procedure and the risk of swapping- and laboratory-mistakes.	270	Y	26	4	NA	Patient questionnaire
2. The chances of success, pregnancy after SET and DET	300	Y	33	2	NA	Patient questionnaire
3. the chances of congenital diseases/malformations, ectopic pregnancy, multiple pregnancy and spontaneous abortion and possible unknown long-term risks.	256	Y	32	5	NA	Patient questionnaire
4. The emotional aspects of the treatment, how patients can communicate that they are in need of extra emotional support and the existence of the patient association for infertility to share experiences and gain information from the patient's perspective.	280	Y	52	3	Y	Patient questionnaire
5. How hospital care during treatment is organised and which caretakers are involved in the treatment procedure, how patients should contact the fertility team in case of problems or complaints and how the fertility team can be contacted, especially outside daytime-hours.	264	Y	39	4	Y	Patient questionnaire

Quality indicators of the 'Embryo Act'.

[a] case mix correction is necessary for 'duration of infertility'

[b] case mix correction is necessary for 'duration of infertility', 'female age' and 'type of infertility for the couple'

[d] this indicator applies to clinics with an IVF laboratory, not patients

N=no, Y=yes, NA=not applicable, SET=single embryo transfer, DET=double embryo transfer

Table IIIf. Quality criteria of guideline-based indicators for clinical fertility care: an assessment in 16 Dutch fertility clinics.

Tubal Pathology	Applicability ($n_{patients}$)	Improvement potential	Discriminatory capacity (%)	Complexity ($n_{variables}$)	Case mix stability	Data source
In case of IVF because of inoperable tubapathology, salpingectomy shoud be performed when bilateral hydrosalpinges are visible by ultrasound.	0	NA	NA	4	NA	Medical record
In case of tubal pathology, both diagnostic laparoscopy and complementary examination of the endosalpinx (hysterosalpingogram or salpingo-scopy) should be performed to decide in favour of tubal surgery.	1	NA	NA	3	NA	Medical record
The indication for tubal surgery in a patient should be made by the future surgeon himself (i.e. videomaterial or by performing a diagnostic laparoscopy).	0	NA	NA	2	NA	Medical record
Except for refertilisation, the female age limit for tubal surgery should be 40 years.	4	NA	NA	3	NA	Medical record

Quality indicators of the guideline 'Tubal pathology'.

Table IIIg. Quality criteria of guideline-based indicators for clinical fertility care: an assessment in 16 Dutch fertility clinics.

Endometriosis	Applicability ($n_{patients}$)	Improvement potential	Discriminatory capacity (%)	Complexity ($n_{variables}$)	Case mix stability	Data source
Prior to laparoscopy intended to diagnose endometriosis, there should be done at least an history-taking with a suspect history, a vaginal and speculum examination.	9	NA	NA	5	NA	Medical record
Women with peritoneal endometriosis and infertility should not be given hormonal treatment to improve fecundicity.	0	NA	NA	3	NA	Medical record
In case of unexplained fertility and peritoneal endometriosis, IUI with ovarian stimulation should be given as primary treatment to improve fecundity.	0	NA	NA	5	NA	Medical record

Quality indicators of the guideline 'Endometriosis'.

[a] case mix correction is necessary for 'duration of fertility'; [b] case mix correction is necessary for 'duration of infertility', 'female age' and 'type of infertility for the couple'; N=no, Y=yes, NA= not applicable

Table IIIh. Quality criteria of guideline-based indicators for clinical fertility care: an assessment in 16 Dutch fertility clinics.

Ovarian Hyper Stimulation Syndrome (OHSS)	Applicability ($n_{patients}$)	Improvement potential	Discriminatory capacity (%)	Complexity ($n_{variables}$)	Case mix stability	Data source
In case of ovulation-induction with gonadotropins, no ovulatory hCG dosis should be given and protected intercourse should be advised in case of: 1. the presence of 3 or more dominant follicles ≥ 18mm or 2. the presence of 5 or more follicles ≥ 15mm and/or 3. A raise in serum-estradiol levels >3,0 nmol/l.	3	NA	NA	6	NA	Medical record
In case of OHSS with hemoconcentration ≥ 45%, the patient should be admitted to the hospital.	2	NA	NA	2	NA	Medical record
In case of severe OHSS and hospital admission, patients should be given trombosis-profylaxis (lowmolecularweight heparin).	2	NA	NA	4	NA	Medical record

Quality indicators of the guideline 'Ovarian Hyper Stimulation Syndrome'.

Table IIIi. Quality criteria of guideline-based indicators for clinical fertility care: an assessment in 16 Dutch fertility clinics.

Premature Ovarian Failure	Applicability ($n_{patients}$)	Improvement potential	Discriminatory capacity (%)	Complexity ($n_{variables}$)	Case mix stability	Data source
Patients with premature ovarian failure, should not be offered any treatment to pursue pregancy, except for oocyte donation.	0	NA	NA	1	NA	Medical record
In case premature ovarian failure is diagnosed, the following should be discussed with the patient: 1. Possibilities to receive psychological support, oocyte-donation, 2. the probabilities of a spontaneous conception, the option of assessing karyotype *(Premature Ovarian Failure, Fragile X, Auto-immune diseases)* 3. the (dis-)advantages of hormone replacement therapy.	0	NA	NA	5	NA	Patient questionnaire
In every patient with premature ovarian failure who chooses *not* to have hormonal suppletion therapy, a bone-densitrometry should be performed.	0	NA	NA	3	NA	Medical record

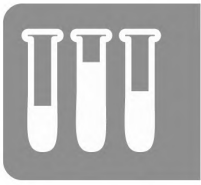
Quality indicators of the guideline 'Premature ovarian failure'.

[a]case mix correction is necessary for 'duration of infertility'; [b]case mix correction is necessary for 'duration of infertility', 'female age' and 'type of infertility for the couple'; N=no, Y=yes, NA= not applicable

References

1. McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, DeCristofaro A, Kerr EA. The quality of health care delivered to adults in the United States. N Engl J Med 2003;348:2635–45.
2. McGlynn EA. Quality assessment of reproductive health services. West J Med 1995;163:19–27.
3. Schuster MA, McGlynn EA, Brook RH. How good is the quality of health care in the United States? Milbank Q 1998;76:517–563, 509.
4. Grol R. Successes and failures in the implementation of evidence-based guidelines for clinical practice. Med Care 2001;39:II46–II54.
5. Greil AL. Infertility and psychological distress: a critical review of the literature. Soc Sci Med 1997;45:1679–704.
6. Kallen B, Finnstrom O, Nygren KG, Otterblad OP, Wennerholm UB. In vitro fertilisation in Sweden: obstetric characteristics, maternal morbidity and mortality. BJOG 2005;112:1529–35.
7. Ombelet W, Martens G, De SP, Gerris J, Bosmans E, Ruyssinck G, Defoort P, Molenberghs G, Gyselaers W. Perinatal outcome of 12,021 singleton and 3108 twin births after non-IVF-assisted reproduction: a cohort study. Hum Reprod 2006;21:1025–32.
8. Cousineau TM, Domar AD. Psychological impact of infertility. Best Pract Res Clin Obstet Gynaecol 2007;21:293–308.
9. Schieve LA, Cohen B, Nannini A, Ferre C, Reynolds MA, Zhang Z, Jeng G, Macaluso M, Wright VC. A population-based study of maternal and perinatal outcomes associated with assisted reproductive technology in Massachusetts. Matern Child Health J 2007;11:517–25.
10. Grimshaw JM, Thomas RE, Maclennan G, Fraser C, Ramsay CR, Vale L, Whitty P, Eccles MP, Matowe L, Shirran L et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. Health Technol Assess 2004;8:iii-72.
11. Grol R. Personal paper. Beliefs and evidence in changing clinical practice. BMJ 1997;315:418–21.
12. Grol R, Baker R, Moss F. Quality improvement research: understanding the science of change in health care. Qual Saf Health Care 2002;11:110–1.
13. Mourad SM, Hermens RP, Nelen WL, Braat DD, Grol RP, Kremer JA. Guideline-based development of quality indicators for subfertility care. Hum Reprod 2007;22:2665–72.
14. Dickey RP, Sartor BM, Pyrzak R. What is the most relevant standard of success in assisted reproduction?: no single outcome measure is satisfactory when evaluating success in assisted reproduction; both twin births and singleton births should be counted as successes. Hum Reprod 2004;19:783–7.
15. Schieve LA, Reynolds MA. What is the most relevant standard of success in assisted reproduction?: challenges in measuring and reporting success rates for assisted reproductive technology treatments: what is optimal? Hum Reprod 2004;19:778–82.
16. Palmer RH. Process-based measures of quality: the need for detailed clinical data in large health care databases. Ann Intern Med 1997;127:733–8.
17. Eddy DM. Performance measurement: problems and solutions. Health Aff (Millwood) 1998;17:7–25.
18. Brook RH, McGlynn EA, Shekelle PG. Defining and measuring quality of care: a perspective from US researchers. Int J Qual Health Care 2000;12:281–95.
19. Mant J. Process versus outcome indicators in the assessment of quality of health care. Int J Qual Health Care 2001;13:475–80.
20. Rubin HR, Pronovost P, Diette GB. The advantages and disadvantages of process-based measures of health care quality. Int J Qual Health Care 2001;13:469–74.
21. Nelen WL, Hermens RP, Mourad SM, Haagen EC, Grol RP, Kremer JA. Monitoring reproductive health in Europe: what are the best indicators of reproductive health? A need for evidence-based quality indicators of reproductive health care. Hum Reprod 2007;22:916–8.
22. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34–42.
23. Rubin HR, Pronovost P, Diette GB. From a process of care to a measure: the development and testing of a quality indicator. Int J Qual Health Care 2001;13:489–96.
24. Campbell SM, Braspenning J, Hutchinson A, Marshall MN. Research methods used in developing and applying quality indicators in primary care. BMJ 2003;326:816–9.

25. Schouten JA, Hulscher ME, Wollersheim H, Braspenning J, Kullberg BJ, van der Meer JW, Grol RP. Quality of antibiotic use for lower respiratory tract infections at hospitals: (how) can we measure it? Clin Infect Dis 2005;41:450–60.
26. Hermens RP, Ouwens MM, Vonk-Okhuijsen SY, van der WY, Tjan-Heijnen VC, van den Broek LD, Ho VK, Janssen-Heijnen ML, Groen HJ, Grol RP, et al. Development of quality indicators for diagnosis and treatment of patients with non-small cell lung cancer: a first step toward implementing a multidisciplinary, evidence-based guideline. Lung Cancer 2006;54:117–24.
27. Ouwens MM, Marres HA, Hermens RR, Hulscher MM, van den Hoogen FJ, Grol RP, Wollersheim HC. Quality of integrated care for patients with head and neck cancer: development and measurement of clinical indicators. Head Neck 2007;29:378–86.
28. Streiner D, Norman G. Health Measurement Scales, 2nd edn. Oxford, UK: Oxford University Press, 1994.
29. Grol R, Wensing M, Eccles M. Improving Patient Care. The Implementation of Change in Clinical Practice, 1st edn. London, UK: Elsevier Limited, 2005.
30. Wollersheim H, Hermens R, Hulscher M, Braspenning J, Ouwens M, Schouten J, Marres H, Dijkstra R, Grol R. Clinical indicators: development and applications. Neth J Med 2007;65:15–22.
31. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.
32. Reading AE. Psychological preparation for surgery: patient recall of information. J Psychosom Res 1981;25:57–62.
33. Kessels RP. Patients' memory for medical information. J R Soc Med 2003;96:219–22.
34. Haagen EC, Nelen WL, Hermens RP, Braat DD, Grol RP, Kremer JA. Barriers to physician adherence to a subfertility guideline. Hum Reprod 2005;20:3301–6.
35. Schmidt L. Infertile couples' assessment of infertility treatment. Acta Obstet Gynecol Scand 1998;77:649–53.
36. Souter VL, Penney G, Hopton JL, Templeton AA. Patient satisfaction with the management of infertility. Hum Reprod 1998;13:1831–6.
37. Malin M, Hemmink E, Raikkonen O, Sihvo S, Perala ML. What do women want? Women's experiences of infertility treatment. Soc Sci Med 2001;53:123–33.
38. Schmidt L, Holstein BE, Boivin J, Tjornhoj-Thomsen T, Blaabjerg J, Hald F, Rasmussen PE, Nyboe AA. High ratings of satisfaction with fertility treatment are common: findings from the Copenhagen Multi-centre Psychosocial Infertility (COMPI) Research Programme. Hum Reprod 2003;18:2638–46.
39. Goverde AJ, McDonnell J, Vermeiden JP, Schats R, Rutten FF, Schoemaker J. Intrauterine insemination or in-vitro fertilisation in idiopathic subfertility and male subfertility: a randomised trial and cost-effectiveness analysis. Lancet 2000;355:13–8.
40. Evers JL. Female subfertility. Lancet 2002;360:151–9.
41. Gagliardi AR, Fung MF, Langer B, Stern H, Brown AD. Development of ovarian cancer surgery quality indicators using a modified Delphi approach. Gynecol Oncol 2005;97:446–456.
42. Pronovost PJ, Miller M, Wachter RM. The GAAP in quality measurement and reporting. JAMA 2007;298:1800–2.



Chapter 5

Information provision in fertility care: a call for improvement

S.M. Mourad
R.P.M.G. Hermens
T. Cox-Witbraad
R.P.T.M. Grol
W.L.D.M. Nelen
J.A.M. Kremer

Human Reproduction 2009; 1(1): 1–7.

Abstract

Background: Adequate information provision is a crucial dimension of high-quality fertility care. Clinical practice guidelines containing consensus-based recommendations may standardize practice between settings. This study was designed for three purposes: (i) to assess actual adherence to recommendations on information provision, (ii) to measure patient satisfaction with current practice and (iii) to analyse how variation in adherence relates to the characteristics of patients and clinics.

Methods: All recommendations concerning patient information were extracted from 10 national fertility guidelines and edited into a patient questionnaire. Additional questions concerning patient satisfaction and potential determinants of information provision at patient level were included. A total of 2698 couples from 16 clinics were invited to participate. A professional's questionnaire was sent to all gynaecologists to gather potential determinants at clinic level. Multilevel regression analysis was performed to identify the determinants of information provision.

Results: A total of 1499 couples (56%) participated. The percentage of couples who reported to have received complete information varied between recommendations from 10 to 96% (mean 57%). Overall, 94% of couples were satisfied with fertility services. The use of checklists for information provision, the presence of obstetrics/gynaecology residents and specialized nursing personnel, and higher patient anxiety scores were significantly associated ($P < 0.05$) with higher levels of information received.

Conclusions: Despite the possibility of recall bias in questionnaire studies and observed high patient satisfaction with fertility services, we conclude that information provision for infertile couples is currently poor and in need of improvement. This could easily be procured by, for example, the use of information checklists.

Introduction

Adequate information provision is a crucial dimension of patient-centred and high-quality care. Good communication and full comprehension of information have been reported as important prerequisites for obtaining informed consent and achieving patient satisfaction, hence improving compliance with doctor's advice and treatment outcomes.[1-3] Providing patient information should therefore not merely be intended to prepare patients on the actual course of their treatment, but more importantly, it should enable them to participate actively and be well informed in the agreed treatment plan.[4-8] The importance of information provision is acknowledged by diverse international organizations, such as the European Patient Federation and the World Health Organization (WHO), who make a clear statement on the topic within the 'declaration on the promotion of patients' rights in Europe':

> *'Patients have the right to be fully informed about their health status, including the medical facts about their condition; about the proposed medical procedures, together with the potential risks and benefits of each procedure; about alternatives to the proposed procedures, including the effect of non-treatment; and about the diagnosis, prognosis and progress of treatment'.*[4]

In fertility care, a patient's decision on commencing treatment could signify considering elective surgery, assisted reproductive technologies or even the use of donor gametes. These are weighty choices and it is therefore not surprising that infertile couples rate information provision as one of the most important aspects of good clinical care.[9-12]

Unfortunately, there is no evidence for which information exactly should be provided to different patient groups. Based on the cited universal WHO statement, however, consensus-based recommendations on information provision can be formulated per practice topic and included in wide-spread clinical practice guidelines. For fertility care, such recommendations are, for instance, already included in selected guidelines of the European Society of Human Reproduction and Endocrinology, the National Institute of Clinical Excellence and the Dutch Society of Obstetrics and Gynaecology (NVOG). Recommendations concern, for example, the provision of general treatment information, but also of information regarding risks and complications of treatment. Clinical practice guidelines can thus facilitate individual physicians to standardize information provision in the consulting room.

The question remains, however, whether these guideline recommendations are actually followed, as guidelines are not self-implementing.[13] The objectives of this study were therefore first to assess actual adherence to such guideline recommendations on information provision. Second, we measured patient satisfaction with this current practice, to evaluate its correspondence with the level of information provision. Since understanding current practice by the recognition of potential determinants is an important first step towards achieving optimal care[14], for our third objective we analysed the extent to which variation in adherence was related to certain patient or clinic characteristics.

Materials and Methods

Setting

We conducted a cross-sectional postal survey study, consisting of a questionnaire and two reminder rounds. The study was approved for all clinics by the 'Regional Review Board for Human Research (CMO) Arnhem-Nijmegen (CMO no. 2004/193)' and is part of a larger research project on which we previously published.[15] The NVOG issued nine national fertility guidelines to facilitate professionals in providing effective and evidence-based care. Guidelines encompass both diagnostics and treatment of fertility problems. The model protocol of the Dutch Embryo Act is in daily practice also used as a complementary national guideline on in vitro fertilization (IVF). These 10 documents describe the minimal degree of patient information that should be given prior to or during fertility treatment. All recommendations concerning patient information were extracted from these national guidelines by four of the authors (S.M., W.N., R.H. and J.K.). Selected recommendations (n=18) comprised several domains: general information about treatment, risks of treatment, possible complications, lifestyle change, psychosocial and medical follow-up. Because data on information provision is hard to extract reliably from medical records[16], all 18 recommendations were edited into a patient questionnaire.

Study population

We included a representative Dutch patient group, visiting 16 participating clinics for diagnosis or treatment of infertility. These clinics vary in size, offer different treatment options (e.g. including IVF/intra-cytoplasmic sperm injection (ICSI) or not) and are both teaching and non-teaching, primary, secondary and tertiary referral clinics. Potential participating couples were selected by means of each clinic's diagnosis treatment combination code registration database; in this national financial registration, fertility patients are identified with a specific fertility-code (F-code). An F-code could mean either an initial assessment of fertility, diagnostics for fertility, fertility treatment (e.g. surgery, ovulation induction, intra-uterine insemination or IVF) or only counselling. Couples were apt for inclusion if they had an F-code anytime in April, May or June 2005. In each clinic, a random group of infertile couples was invited to take part in the study and was sent an informed consent form and a questionnaire. The patient sample size was chosen according to the clinic size (50, 150 or 550 eligible patients); a total of 2698 couples were invited. Couples who completed both the questionnaire and informed consent forms were included for analysis. Couples who did not have enough knowledge of the Dutch language to fill out the questionnaire or turned out not to have visited the clinic in the requested period were excluded.

Questionnaire development

The patient questionnaire was constructed to address the three objectives of this study:

Part 1: actual adherence to recommendations on information provision

Questions were carefully formulated to detect a couple's direct experience of care; they were asked whether they had received specific information from their own care-providers (as opposed to a formulation directed at testing their knowledge of specific topics). For example: 'did your clinic provide you with information about the risk of

ectopic pregnancy before you started an IVF/ICSI treatment?' Questions were composed of four closed response categories, 'yes', 'no', 'cannot remember'and 'irrelevant'.

Part 2: patients' satisfaction with information provision
Satisfaction ratings of the female partner were measured by a Dutch translation of Souter's 'questionnaire on patient satisfaction with the management of infertility'.[10]

Part 3: determinants of information provision
We searched for potential determinants of information provision from the literature regarding fertility as well as other fields of health care.[14,17-22]

These potential determinants were hypothesized to be found in three groups, which are shown in Supplementary Table 1a:
(1) Patient characteristics (at individual and couple level, e.g. female age, couple's education level, female's anxiety for treatment)
(2) Clinic's organization of fertility care (e.g. clinic size, IVF/non-IVF facilities, presence of trained fertility nurses).
(3) Clinic's organization of information provision (e.g. availability of a lifestyle change programme, organization of an informative meeting, use of information checklists).

Questions regarding demographic characteristics and potential determinants based on patient characteristics (both at the level of individual partners as well as the couple) were added to the patient questionnaire. The female partner was asked to fill out additional questions on anxiety and depression. Anxiety was measured by a 10 item short version of the state trait anxiety index (STAI)[23,24] and n=12 additional infertility-related anxiety items, e.g. 'anxiety for treatment outcome, both on a four-point scale ("almost never", "sometimes", "frequently" and "almost always")'. Depression was measured by the Beck Depression Index for primary care (BDI-PC) that uses a four-point scale for varying utterances.[25]
When completing the questionnaire, patient couples were asked to describe their experiences during the study period (1 January 2005 to 1 July 2005). Because a fertility problem affects both partners who are frequently seen in a joint consultation, couples were asked to fill out part I of the questionnaire (regarding the information recommendations) preferably together. The patient questionnaire was piloted in a group of 30 infertile couples recruited through the website of the Dutch Patients' Association for Infertility 'Freya'. This pilot led to minor adjustments in formulation of some questions before the questionnaire was used in the study group.
A professional questionnaire was composed of questions regarding potential determinants based on a clinic's organization of fertility care and organization of information provision, and sent to a gynaecologist of each of the 16 participating clinics.

Statistical analysis

We used the Statistical Package for the Social Sciences (SPSS 14.0 for Windows, SPSS Inc., Chicago, IL, USA) for most analyses. If a single guideline-derived recommendation encompassed several topics, it was discussed in the study group (S.M., W.N., J.K. and R.H.) until consensus was reached. If considered appropriate, it was subdivided for further analysis. This procedure resulted in 28 recommendations for analysis. Descriptive analysis was performed to assess frequencies of adherence to the recommendations. If a recommendation was applicable to less than 10 patients, it was excluded from further analysis.

We applied univariate analysis (cross tabulations, Wilcoxon's rank sum test, independent samples t-test and X^2 test) to examine the associations between several patient's or clinic's characteristics (independent variables) on the one hand, and adherence to information recommendations (dependent variables) on the other hand. Before applying this analysis, a confirmatory factor analysis and reliability assessment was performed for the questions of the STAI (Cronbach's α=0.91) and BDI-PC (Cronbach's α=0.82); both showed good internal consistency within our study population and thus the sumscores 'depression' and 'state anxiety' were carried forward as potential determinants. Exploratory factor analysis was performed on the additional questions regarding anxiety to reduce the number of potential determinants. Regarding this, the following two factors were identified: anxiety for treatment (Cronbach's α=0.73) and anxiety for treatment outcome (Cronbach's α=0.76). Remaining anxiety items that did not belong to a factor were 'anxiety for financial consequences' and 'anxiety for relationship with partner'; these were also treated as potential determinants.

All independent variables were subsequently analysed for colinearity. If a correlation between two independent variables was detected (correlation coefficient > 0.4), the most relevant candidate variable with respect to content was carried forward. All independent variables, which were found to be univariately significantly associated with adherence to the information recommendations ($P < 0.10$) and showed enough variation between the different clinics, were included in a multilevel stepwise logistic regression analysis to explain difference in adherence. If variation between clinics was nil, a regular multivariate regression analysis was performed. For the multilevel analysis, a random coefficient model was composed using two levels (clinic and patient) in a Glimmix procedure in SAS (SAS for Windows V8.2). Significance for both multivariate as well as multilevel analysis was set at $P < 0.05$.

Results

Response

A total of 1499 couples (=56%) completed the questionnaire and gave their consent to join the study. In only 24%, the couple did fill out the questionnaire together; in 76% it was completed exclusively or mainly by the woman (i.e. sporadically asked her partner for an answer). Socio-demographic characteristics of the participating couples are shown in Table I.[15] Mean female age was 32.8 years and male age 35.1 years. Only 1.9% (n=27) of couples were of non-Dutch origin (i.e. both partners were non-Dutch). Of the participating couples, 73% suffered from primary infertility and the median duration of infertility was 38 months. A total of 53% had a high education level (more

than secondary school) and 94% had a more than modal income (per household > 1760 euro/month gross).

Actual adherence to information provision recommendations

Information recommendations on 'premature ovarian failure' and 'cancellation criteria for intra-uterine insemination' were excluded from analysis because ,10 patients were eligible for analysis. The percentage of couples who reported to have received complete information for one of the recommendations ranged from 10 to 96% (mean 57%) and is shown in Supplementary Table 2a. The information recommendation that scored lowest, with 10%, concerns the risk and symptoms of ectopic pregnancy after tubal surgery, closely followed by the recommendation 'to discuss risks of an IVF/ICSI treatment prior to actual treatment' with 14%. The best scoring recommendations were 'having an evaluative consultation when IVF/ICSI treatment is terminated' (96%) and 'to discuss the assessment of tubal patency both pro- and retrospectively' (95%).
Regarding the different content domains, the information concerning complications (e.g. ovarian hyper stimulation syndrome, OHSS) was received in 32% of appropriate cases, information concerning risks (e.g. ectopic pregnancy) in 41% of cases, followed by lifestyle advise (e.g. weight loss) in 46%. Higher scores of 72% were found for both general information (e.g. accessibility of the clinic within and outside office hours) and information concerning additional emotional or psychological support (e.g. contact information of the patient association).

Patient satisfaction with information provision

In total, 35% of the couples mentioned information provision as the most important aspect of care compared with: waiting time in clinics, doctor's attitude, the way investigations are done and emotional support. However, 26% of couples wished to have had more written information, whereas 68% reported to have received any written information on diagnostics, background and treatment of their infertility. Overall, 94% of the couples were satisfied or very satisfied with their current fertility services. Couples who achieved pregnancy in the study period were significantly more satisfied (P=0.000) than those who did not. Satisfaction was however not significantly associated (P= 0.515) with pregnancy outcome.

Table I. Baseline characteristics of the participating couples (n=1499)

Characteristics	% of couples
Mean age in years (SD)	
Female	32.8 yrs (4.1)
Male	35.1 yrs (5.0)
Ethnic background[a]	
Dutch	98.1
Non-Dutch	1.9
Gross monthly family income (euros)[b]	
<1100	1.6
1100-1760	4.6
1760-2750	22.2
>2750	71.6
Education level per couple[c]	
Low	6.4
Intermediate	40.7
High	52.9
Type of subfertility[d]	
Primary	72.8
Secondary	27.2
Median duration of subfertility in months (SD)[e]	38.4 (22.5)

[a] Ethnic background of the couples was determined by the origin of both partners: Dutch = one or both partners are of Dutch origin; non-Dutch = both partners are not of Dutch origin.

[b] Gross monthly family income was categorized according to social security standards in 2005 and modal income in euros:<1100 less than Dutch modal income; 1100-1760 Dutch modal income; 1760-2750 up to 1,5 times Dutch modal income; >2750 more than twice Dutch modal income.

[c] Education level of the couples was determined by the highest education level of both partners: low = primary or lower vocational education; intermediate = secondary or intermediate vocational education; high = higher professional education or university.

[d] Type of subfertility was determined for the couple.

[e] Duration of subfertility was defined as the period between the start of regular unprotected sexual intercourse and 1 January 2005, the beginning of the study period.

Determinants of information provision

As hypothesized, we found univariate significant associations within the three groups of determinants: patient's characteristics, clinic's organization of fertility care and clinic's organization of information provision (data not shown). Of the 28 information recommendations, 7 showed sufficient variation between the clinics to be analysed subsequently in a multilevel procedure. The results of these multilevel analyses are shown in Table II. At patient level, there was a positive association between, on the one hand, high education level (P=0.0064) and high treatment-related anxiety scores (P=0.0275, P=0.0050, P=0.0289 and P=0.0070), and on the other hand, a higher level of received information concerning 'lifestyle advice' about weight, alcohol and drug use and 'emotional consequences of treatment'. At the level of a clinic's organization of fertility care, the presence of obstetrics/gynaecology residents in the fertility department positively influenced 'information provision on lifestyle advice concerning weight' (P=0.0017) and 'emotional consequences of treatment' (P=0.0375). Moreover, the presence of trained fertility nurses positively influenced 'information provision on prevention of OHSS' (P=0.0390) and 'lifestyle advice concerning tobacco use' (P=0.0023). A clinic's higher number of consultations per year was positively associated with information concerning alcohol use (P=0.0277). At the level of a clinic's organization of information provision, the use of checklists for information provision

was positively associated with a higher level of information provision on 'prevention of OHSS' (P=0.0294 and P=0.0024) and 'lifestyle advice on alcohol use' (P=0.0241).

Discussion

In this study, we observed that for Dutch patients information provision according to national fertility guideline recommendations is currently poor and in need of improvement. Patient couples are deprived of essential information on diagnostics, causes of their condition and treatment risks. Determinants of high information provision are, e.g. patients' high education level, the presence of trained fertility nurses and the use of information checklists for professionals.
However, the majority of patients view the information they receive as sufficient, and patient satisfaction with treatment is high.

Adherence to information recommendations

We showed that the percentage of couples who received complete information varied widely per guideline recommendation. It is especially alarming that the majority of the patient couples are deprived of essential information concerning complications and risks of treatment; only 32 and 43% of couples, respectively, actually received this complete information. It is thus disputable whether these couples have really been in the position to make informed decisions before starting treatment. In general, an observed lack of information can be attributed to two main factors, i.e. limited information provision by healthcare professionals and defective memorization or comprehension by patients.[26] If the former is the main problem, professionals should be confronted with their inadequate performance to enable them to improve. However if the latter is the main problem, professionals should also be alert to recognize when information is not sufficiently understood. Furthermore, providing information on complications and risks of fertility treatment can be complicated by the fact that couples perceive treatment merely as a positive thing, i.e. a solution for their problem. They could therefore ignore warnings and downplay any negative or frightening information about the much desired treatment.
In general, infertile patients comprise a relatively young, actively participating and, in our sample, even well-educated patient group.
Several studies in oncology and primary care showed that younger patients and patients striving for active involvement in their treatment have higher intrinsic needs for information; this makes our observed lacunas in information provision, particularly alarming.[21,27] The fact that clinicians incorrectly associate a younger patient age with better understanding of information or with better abilities to gather information themselves, might be a serious pitfall in daily practice.
Healthcare professionals should therefore be attentive to patient's preferences and perspectives and should ensure that information, particularly on patient safety, is actually taken in and fully understood.

Satisfaction

To deliver patient-centred healthcare, professionals should put effort into collecting, as well as acting upon, patients' preferences on a regular basis. The results of our survey showed that the vast majority (94%) of participating couples were satisfied with

their fertility services, with even higher ratings when pregnancy was achieved. Likewise high ratings of patient satisfaction are known from the literature.[10,12,28,29] However, such high ratings may mask still existing shortcomings in actual care, especially as ratings were shown to be influenced by desired outcome (i.e. pregnancy). For example, a patient who is unaware that information provision is incomplete has no reason to be critical and is therefore likely to be content with the information he or she received. It is important to realize that a patient satisfaction assessment is only an indirect, and therefore insufficient, method to monitor current practice; it should be completed with more care-related and preferably evidence-based evaluations of actual performance, such as the current study.

Determinants of information provision

In each of the three hypothesized domains (Supplementary Table 1a), determinants were found to be significantly related to information provision levels after multilevel analysis. Within the domain 'clinic's organization of fertility care', the presence of specialized fertility nurses and residents was found to be associated with better information provision on selected topics, as was the use of information checklists from the domain 'clinic's organization of information provision'.

This probably reflects the typically more systematic working methods of both groups in comparison with the more autonomous and routine practice of, for example, established gynaecologists. It also means that a simple and systematic approach, such as the introduction of information checklists for professionals, can compensate for organizational characteristics such as a clinic's small size, low number of fertility consultations or the lack of specialized nursing personnel.

The observed association between higher levels of information received and higher treatment-related anxiety scores raises the question that a causal connection could exist. Are more anxious patients craving for information and thus better informed, or does an overload of warning information make patients confused and more anxious, even affecting memorization? This controversy is previously described in the literature[30-34], and further research is needed to try to unravel the underlying mechanisms. In the meanwhile, special attention should be paid to anxiety levels of infertile patients.

Table II. Significant determinants of information provision after multi-level analysis

Recommendation	Determinant of high information provision	P-value
In case of actual mild OHSS, patients should be given the following information to monitor signs of OHSS	Use of checklists for information provision	0.0294
Increase fluid intake		
Check colour of urine		
Daily monitoring of body weight in the morning		
In case of actual mild OHSS and an increase in body weight of >1 kg/day, patients should be advised to contact the clinic for ultrasound and laboratory tests	Presence of trained fertility nurse	0.0390
	Use of checklists for information provision	0.0024
During an intake prior to IVF/ICSI-treatment, the following should be discussed:	High level of anxiety for relationship with partner	0.0275
Emotional aspects of the treatment	Presence of Ob/Gyn residents	0.0375
How patients can communicate that they are in need of extra emotional support		
The existence of the subfertility patient association 'Freya' for the possibility to share experiences and gain information from the patients perspective		
Patients with anovulation and overweight should, with regard to their fertility treatment and overall health, be informed of the importance of weight reduction by means of life-style changes[a].	High level of anxiety for treatment outcome	0.0050
Lifestyle advice regarding tobacco, should be part of the counselling regarding pregnancy probabilities[a].	Presence of trained fertility nurses	0.0023
	High level of anxiety for treatment	0.0289
Lifestyle advice regarding alcohol should be part of the counselling regarding pregnancy probabilities[a].	Couple's high education level	0.0064
	High number of fertility consultations/year	0.0277
	Use of checklists for information provision	0.0040
Lifestyle advice regarding a healthy weight should be part of the counselling regarding pregnancy probabilities[a].	High level of State anxiety	0.0289
	High level of anxiety for treatment outcome	0.0070
	Presence of Ob/Gyn residents	0.0017

Ob/Gyn=Obstetrics and Gynaecology; OHSS=ovarian hyper stimulation syndrome. Significance was set at $p<0.05$.

[a] patients were only included for analysis of this recommendation when they respectively used alcohol, tobacco or had a BMI <20 or >25 kg/m^2

Advantages and disadvantages of the study

An advantage of this study is that the investigated content domains of information in this study are universal for fertility care and not unique to the Dutch situation, which makes the reported results also of considerable interest to other countries. It shows that the existence of best practice guidelines does not automatically coincide with actual best practice. Acknowledging this gap between current and best practice will be the first important step towards optimal information provision.

Adequate ways should be found to guide clinicians in taking the next steps. The attitude towards patient participation in guideline development could, for example, precipitate the development of patient information as a regular addendum to each newly developed guideline, as suggested before by Coulter et al.[5] .

This study also has some limitations. First, recall bias should be considered whenever analysing patient questionnaires. A 100% adherence score per recommendation might not be feasible due to incomplete patient recall.[35-37] However, the large differences in adherence scores between the recommendations as well as between the participating hospitals (data not shown) illustrate that recall bias might not be the most important nor the sole explanation for the reported low adherence scores, as each recommendation is a priori at comparable risk for this recall bias. We tried to minimize the effect of possible recall bias by choosing a study design with a questionnaire aimed at couples. It is certainly possible that one of the partners cannot correctly recall the information provided, but the other partner may be able to compensate this lack of recall. We hypothesized that addressing the couple instead of the individual will give us a more reliable representation of actual practice. Conversely, the questions concerning anxiety, depression and satisfaction were exclusively aimed at the female partner. The literature, however, shows that gender differences can be present for these scales.[12,38,39] Future research would therefore benefit from a design directed at distributing separate questionnaires to the individual partners. Such an adjusted design would make it possible to compare scores from both partners, thus shedding more light on the role of the male partner as well as couple dynamics regarding the issue of patient information in fertility care.

Secondly, the use of an extensive written questionnaire may have caused a sample bias; Non-Dutch or low-literacy couples are less apt to participate because of insufficient Dutch language skills. It is important to realize that the level of received information in these patient groups will be even less than our reported findings, i.e. our results could be an under-estimation of reality. Both groups already tend to ask fewer questions during consultation and are therefore at risk to understand less of the provided information.[21,22,40] Supplemental written information to consolidate oral information is known to improve recall[41,42], but has no effect in these distinctive patient groups. Further research should therefore aim at identifying, acknowledging and acting upon specific informational needs of certain patient subgroups. This could result in more tailored information provision, e.g. using pictograms for low-literacy patients[43,44], or providing translated information material for non-native speakers.[45]

Thirdly, some patient characteristics known from the literature to be related to transmission of information (e.g. ethnicity, social class, age and income)[17] were not found to be significant in our analysis. A possible explanation is that our sample of infertile couples was a relatively homogenous group compared with, for example,

patient groups in oncology or primary care; the age-range was relatively small and ethnicity, social class and income showed a skewed distribution. Sufficient data variation between the clinics is necessary to analyse all potential determinants in a multilevel procedure. Future research in a larger multicentre study should be performed to overcome this lack of inter-patient, inter-couple and inter-clinic variation. In conclusion, we found that information provision in fertility care in the Netherlands is currently poor. This is despite the effort of professional organizations such as the NVOG and the Dutch Ministry of Health, Welfare and Sport to describe the minimal informational needs of infertile patients in official nationwide guidelines and even an act of law. Nevertheless, we reported high ratings of patient satisfaction, suggesting that such ratings alone are insufficient to assess actual care. Information provision could easily be improved by, for example, the use of information checklists.

Table 1a Potential determinants of information provision

Determinant domain	Level	Determinant
Patient	**Individual**	Female age (years) Male age (years) Alcohol use (yes/no) both partners Tobacco use (yes/no) both partners Drug use (yes/no) both partners State anxiety [a] female partner Anxiety for treatment[a] female partner Anxiety for treatment outcome[a] female partner Anxiety for financial consequences[a] female partner Anxiety for relationship with partner[a] female partner Depression[a] female partner
	Couple	Type of subfertility (primary/secondary) Duration of subfertility (months) Ethnicity (Dutch/non-Dutch) Education level (low, intermediate, high) Gross monthly family income (euros) Attained pregnancy (yes/no)
Clinic: organization of fertility care		Clinic size (new patients/year) Clinic type (1-16) Number of subfertility consultations/year IVF facilities (yes/no) Teaching facility (yes/no) Presence of separate subfertility outpatient clinic (yes/no) Presence of separate subfertility consulting hours (yes/no) Presence of trained subfertility doctors (yes/no) Number of trained subfertility doctors Presence of trained subfertility nurses (yes/no) Number of trained subfertility nurses Presence of Ob/Gyn residents (yes/no) Number of professionals within subfertility team Local fertility protocols available (yes/no) National fertility protocols available (yes/no)
Clinic: organization of information provision		Lifestyle alteration programme available (yes/no) Locally developed patient-information available (yes/no) Nationally developed patient-information available (yes/no) National patient association information available (yes/no) Organization of informative meetings (yes/no) Use of checklists for information provision (yes/no)

[a] Anxiety and Depression were reflected in sumscores, as described in the text.
Ob/Gyn; Obstetrics and Gynaecology

Table 2a Mean adherence to guideline-based recommendations on information provision

Information domain	Adherence (%)	$n_{patients}$	Guideline source
General information	**72**		
Before prescribing Metformin for ovulation induction because of apparent Clomiphene-resistancy, the following should be discussed: • Possible side effects of Metformin • The off-label use of Metformin for ovulation-induction	59	17	Anovulation and childwish
Each step in the initial assessment of fertility should be discussed with the couple both pro- and retrospectivel;			Initial assessment of
• Cycle-analysis	91	337	fertility
• Semenanalysis	85	316	
• Tubal patency assessment	95	159	
All diagnostics within the initial assessment of fertility should aim at providing the couple insight into their chance of spontaneous conception	83	342	Initial assessment of fertility
In case of ovulation induction to improve fecundity and poor ovarian response(<3 follicles), IVF should be advised against.	26	61	Embryo act
After each unsuccessful cycle of IVF/ICSI, the cycle should be evaluated with the couple and a proposal for possible further treatment should be given.	59	294	Embryo act
During an intake prior to IVF/ICSI-treatment, the chances of pregnancy in general and after single and double embryo transfer respectively[b] should be discussed.	86	301	Embryo act
During an intake prior to IVF/ICSI-treatment, the following should be discussed: • how hospital care during treatment is organised and which caretakers are involved in the treatment procedure • how patients should contact the fertility team in case of problems or complaints • how the subfertility team can be contacted, especially outside daytime-hours	68	265	Embryo act
The patient who is candidate for tubal surgery should be informed about the admittance procedure, the surgery itself and the post-operative recovery.	82	55	Tubal pathology
The patient who is candidate for tubal surgery should be informed about the success rate after surgery	68	53	Tubal pathology
The patient who is candidate for tubal surgery should be informed about the time interval between surgery and possible conception	57	49	Tubal pathology
Lifestyle	**46**		
Patients with overweight should, with regard to their fertility-treatment and overall health, be informed of the importance of weight reduction by means of life-style changes.	28	128	Anovulation and childwish
Lifestyle advice regarding tobacco use should be part of the counselling regarding pregnancy-probabilities.[a]	57	114	Initial assessment of fertility
Lifestyle advice regarding alcohol use should be part of the counselling regarding pregnancy-probabilities.[a]	49	192	Initial assessment of fertility
Lifestyle advice regarding drug use should be part of the counselling regarding pregnancy-probabilities.[a]	NA	7	Initial assessment of fertility

Information domain	Adherence (%)	$n_{patients}$	Guideline source
Lifestyle advice regarding a healthy weight should be part of the counselling regarding pregnancy-probabilities.[a]	50	145	Initial assessment of fertility
Complications	**32**		
During an intake prior to IVF/ICSI-treatment, the following risks should be discussed: • Risk of hyperstimulation • Risk of poor response and accompanying consequences • Risk of complications such as infection and bleeding • The laboratory procedure and the risk of swapping- and laboratory-mistakes	14	271	Embryo act
In case of actual mild OHSS, patients should be given the following information to monitor signs of OHSS: • Increase fluid intake • Check colour of urine • Daily monitoring of body weight in the morning	20	150	OHSS
In case of actual mild OHSS and an increase in body weight of >1 kilogram/day, patients should be advised to contact the clinic for ultrasound and lab-tests.	61	150	OHSS
Additional support	**72**		
During an intake prior to IVF/ICSI-treatment, the following should be discussed: • Emotional aspects of the treatment • How patients can communicate that they are in need of extra emotional support • The existence of the subfertility patient association 'Freya' for the possibility to share experiences and gain information from the patient's perspective	48	280	Embryo act
In case IVF/ICSI-treatment is terminated, an evaluative consultation should be held with the couple to determine whether further follow-up or adjuvant emotional support is necessary.	96	47	Embryo act
In case premature ovarian failure is diagnosed, the following should be discussed with the patient: • Possibility to receive psychological support • The possibility of oocyte-donation • The chances of spontaneous conception • Family-history, with regard to the option of assessing karyotype *(premature ovarian failure, Fragile X syndrome, Auto-immune diseases)* • *The (dis-)advantages of hormone replacement therapy*	NA	0	Premature ovarian failure
Risks	**41**		
Before ovulation induction treatment, the patient should be informed about: • Specific side-effects of medication • The need for regular intensive follow-up during treatment • The increased risk of multiple pregnancy • The increased risk of OHSS • The increased risk of spontaneous abortion	18	204	Anovulation and childwish
During ovulation-induction, the patient has to be informed that coitus is prohibited or contraception has to be used in case of 4 or more follicles >12mm.	67	46	Anovulation and childwish

Information domain	Adherence (%)	$n_{patients}$	Guideline source
Intra Uterine Insemination should not be performed in case of more than 3 follicles > 16 mm or more than 5 follicles > 12mm. In both cases, the use of contraceptives should be advised as well.	92	100	Intra uterine insemination
During an intake prior to IVF/ICSI-treatment, the following risks should be discussed: • Chance of congenital diseases/malformations • Chance of ectopic pregnancy • Chance of multiple pregnancy • Chance of spontaneous abortion • Possible unknown long-term risks	16	257	Embryo act
The patient who is candidate for tubal surgery should be informed about the risk and symptoms of extra-uterine gravidity	10	49	Tubal pathology

[a] patients were only included for analysis of this recommendation when they respectively used alcohol, tobacco or had a BMI <20 or >25 kg/m^2

[b] in the Netherlands, a maximum of two embryos are transferred per treatment cycle.

OHSS = ovarian hyperstimulation syndrome;
IVF= in vitro fertilization;
ICSI= intracytoplasmatic sperm injection

References

1. Ley P. Communicating With Patients: Improving Communication, Satisfaction and Compliance. Croom Helm Ltd, 1989.
2. Weinman J. Providing written information for patients: psychological considerations. J R Soc Med 1990;83:303–5.
3. Simpson M, Buckman R, Stewart M, Maguire P, Lipkin M, Novack D, Till J. Doctor–patient communication: the Toronto consensus statement. BMJ 1991;303:1385–7.
4. WHO. A Declaration on the Promotion of Patients' Rights in Europe: Principles of the Rights of Patients in Europe: A Common Framework. Copenhagen, Denmark: WHO Regional Office for Europe, 1994. http://www.who.int/genomics/public/eu_declaration1994.pdf.
5. Coulter A, Entwistle V, Gilbert D. Sharing decisions with patients: is the information good enough? BMJ 1999;318:318–22.
6. Say RE, Thomson R. The importance of patient preferences in treatment decisions—challenges for doctors. BMJ 2003;327:542–5.
7. Epstein RM, Alper BS, Quill TE. Communicating evidence for participatory decision making. JAMA 2004;291:2359–66.
8. EPF. The European Patients' Forum Position Statement on Information to Patients. 2007. http://www.eu-patient.eu/news/attached_documents/EPF_itp_position_fin.pdf.
9. Schmidt L. Infertile couples' assessment of infertility treatment. Acta Obstet Gynecol Scand 1998;77:649–53.
10. Souter VL, Penney G, Hopton JL, Templeton AA. Patient satisfaction with the management of infertility. Hum Reprod 1998;13:1831–36.
11. Malin M, Hemmink E, Raikkonen O, Sihvo S, Perala ML. What do women want? Women's experiences of infertility treatment. Soc Sci Med 2001; 53:123–33.
12. Schmidt L, Holstein BE, Boivin J, Tjornhoj-Thomsen T, Blaabjerg J, Hald F, Rasmussen PE, Nyboe AA. High ratings of satisfaction with fertility treatment are common: findings from the Copenhagen Multi-centre Psychosocial Infertility (COMPI) Research Programme. Hum Reprod 2003;18:2638–46.
13. Grol R. Personal paper. Beliefs and evidence in changing clinical practice. BMJ 1997;315:418–21.
14. Haagen EC, Nelen WL, Hermens RP, Braat DD, Grol RP, Kremer JA. Barriers to physician adherence to a subfertility guideline. Hum Reprod 2005;20:3301–6.
15. Mourad SM, Nelen WL, Hermens RP, Bancsi LF, Braat DD, Zielhuis GA, Grol RP, Kremer JA. Variation in subfertility care measured by guideline-based performance indicators. Hum Reprod 2008; 23:2493–500.
16. McColl A, Roderick P, Smith H, Wilkinson E, Moore M, Exworthy M, Gabbay J. Clinical governance in primary care groups: the feasibility of deriving evidence-based performance indicators. Qual Health Care 2000;9:90–7.
17. Waitzkin H. Doctor–patient communication. Clinical implications of social scientific research. JAMA 1984;252:2441–6.
18. Davis DA, Taylor-Vaisey A. Translating guidelines into practice. A systematic review of theoretic concepts, practical experience and research evidence in the adoption of clinical practice guidelines. CMAJ 1997;157:408–16.
19. Lukoschek P, Fazzari M, Marantz P. Patient and physician factors predict patients' comprehension of health information. Patient Educ Couns 2003;50:201–10.
20. de Boer EJ, den Tonkelaar I, Burger CW, van Leeuwen FE. Validity of self-reported causes of subfertility. Am J Epidemiol 2005;161:978–86.
21. Cegala DJ, Street RL Jr, Clinch CR. The impact of patient participation on physicians' information provision during a primary care medical interview. Health Commun 2007;21:177–85.
22. Katz MG, Jacobson TA, Veledar E, Kripalani S. Patient literacy and question-asking behavior during the medical encounter: a mixed-methods analysis. J Gen Intern Med 2007;22:782–86.
23. Spielberger R, Gorsuch R, Lushene R. STAI Manual for the State-Trait Anxiety Inventory 1970. Palo Alto, CA: Consulting Psychologists, 1970.
24. Van Der Ploeg H, Defares P, Spielberger C. Handleiding bij de Zelf-Beoordelings Vragenlijst ZBV: Een nederlandstalige bewerking Van de Spielberger State-Trait Anxiety Inventory, STAI-DY [Manual of the Self-Evaluation Questionnaire: a Dutch version of the State-Trait Anxiety Inventory]. Lisse, The Netherlands: Swets and Zeitlinger B.V., 1980.

25. Beck AT, Guth D, Steer RA, Ball R. Screening for major depression disorders in medical inpatients with the Beck Depression Inventory for Primary Care. Behav Res Ther 1997;35:785–91.
26. van den Berg H, Langeveld NE. Parental knowledge of fertility in male childhood cancer survivors. Psychooncology 2008;17:287–91.
27. Ankem K. Factors influencing information needs among cancer patients: a meta-analysis. Libr Inf Sci Res 2006;28:7–23.
28. Sabourin S, Wright J, Duchesne C, Belisle S. Are consumers of modern fertility treatments satisfied? Fertil Steril 1991;56:1084–90.
29. Godwin Y. Do they listen? A review of information retained by patients following consent for reduction mammoplasty. Br J Plast Surg 2000; 53:121–5.
30. Wallace LM. Communication variables in the design of pre-surgical preparatory information. Br J Clin Psychol 1986;25:111–8.
31. Elsass P, Eikard B, Junge J, Lykke J, Staun P, Feldt-Rasmussen M. Psychological effect of detailed preanesthetic information. Acta Anaesthesiol Scand 1987;31:579–583.
32. Sorg BA, Whitney P. The effect of trait anxiety and situational stress on working memory capacity. J Res Pers 1992;26:235–41.
33. Kerrigan DD, Thevasagayam RS, Woods TO, Mc Welch I, Thomas WE, Shorthouse AJ, Dennison AR. Who's afraid of informed consent? BMJ 1993;306:298–300.
34. Luck A, Pearson S, Maddern G, Hewett P. Effects of video information on precolonoscopy anxiety and knowledge: a randomised trial. Lancet 1999;354:2032–5.
35. Ley P. Memory for medical information. Br J Soc Clin Psychol 1979; 18:245–55.
36. Kessels RP. Patients' memory for medical information. J R Soc Med 2003; 96:219–22.
37. Hassan E. Recall bias can be a threat to retrospective and prospective research designs. Internet J Epidemiol 2006;3.
38. Verhaak CM, Smeenk JM, van MA, Kremer JA, Kraaimaat FW. A longitudinal, prospective study on emotional adjustment before, during and after consecutive fertility treatment cycles. Hum Reprod 2005; 20:2253–60.
39. Volgsten H, Skoog SA, Ekselius L, Lundkvist O, Sundstrom Poromaa I. Prevalence of psychiatric disorders in infertile women and men undergoing in vitro fertilization treatment. Hum Reprod 2008; 23:2056–63.
40. Timmermans LM, van der Maazen RW, van Spaendonck KP, Leer JW, Kraaimaat FW. Enhancing patient participation by training radiation oncologists. Patient Educ Couns 2006;63:55–63.
41. Patterson C, Teale C. Influence of written information on patients' knowledge of their diagnosis. Age Ageing 1997;26:41–2.
42. Lloyd AJ, Hayes PD, London NJ, Bell PR, Naylor AR. Patients' ability to recall risk associated with treatment options. Lancet 1999;353:645.
43. Houts PS, Bachrach R, Witmer JT, Tringali CA, Bucher JA, Localio RA. Using pictographs to enhance recall of spoken medical instructions. Patient Educ Couns 1998;35:83–8.
44. Houts PS, Witmer JT, Egeth HE, Loscalzo MJ, Zabora JR. Using pictographs to enhance recall of spoken medical instructions II. Patient Educ Couns 2001;43:231–42.
45. van Eijsden M, van der Wal MF, Bonsel GJ. Folic acid knowledge and use in a multi-ethnic pregnancy cohort: the role of language proficiency. BJOG 2006;113:1446–51.

Chapter 6

Determinants of patients' experiences and satisfaction with fertility care

S.M. Mourad
W.L.D.M. Nelen
R.P. Akkermans
J.H.A. Vollebergh
R.P.T.M. Grol
R.P.M.G. Hermens
J.A.M. Kremer

Fertility and Sterility 2009; in press

Abstract

Objective: To assess determinants of patients' experiences and satisfaction with fertility care.

Design: Cross-sectional questionnaire study.

Setting: Sixteen fertility clinics in The Netherlands.

Patients: A total of 1,499 infertile women in The Netherlands who visited a participating clinic in April–June 2005 for diagnostics or treatment.

Main outcome measures: Patients experiences and satisfaction with several aspects of fertility care, and the patient and clinic characteristics that are determinants of those two concepts.

Results: In general, patients' satisfaction with care was high (94%). Waiting times, information provision and emotional support were experienced the least positive aspects of care. Determinants of all care aspects were found to be significant at four different domains: three at patient level, i.e., demographic characteristics, type of received treatment and both general and mental health status, and one at clinic level, i.e., organization of care.

Conclusions: This study provides an increased understanding of the determinants of patients' experiences and satisfaction with fertility care. This enables professionals to tailor their care to specific subgroups of patients and adjust their organization of fertility care where needed. Moreover, the study underlines the need to investigate whether case-mix correction is necessary whenever interpreting patient-surveys on care experiences, because both the patient's and the clinic's characteristics can influence the way that health care delivery is experienced. Demographic background of this regional patient sample was rather homogeneous, which should be taken into account when interpreting results.

Introduction

Traditionally, the quality of health care is rated by technical and physiologic outcome measures such as mortality and morbidity.[1,2] Regarding fertility care, outcomes are thus frequently defined as live birth or complication rate. However, in the last two decades there has been an increasing conviction that patients' opinions have to be included in the evaluation of health care to achieve a more thorough and patient-centered reflection of quality of care.[1,3,4] For this purpose, standardized methods, such as the Picker Patient Experience questionnaire[5] and the Consumer Assessment of Healthcare Providers and Systems (https://www.cahps.ahrq.gov) have been developed to elicit feedback from patients. Such feedback can thus be used as additional outcome measures within quality-of-care assessments, while simultaneously uncovering those aspects of care that are in need of attention or improvement.[6,7] The current disadvantage of such instruments is, however, that they are not always tailored to specific patient groups and only scarcely available for outpatient care, let alone fertility care. This probably explains why most studies within fertility care have sought to develop and validate their own instruments.[8–10]

The question rises whether measuring experiences and satisfaction alone provides us with enough transparency to directly improve care. A logical second step would be to assess which determinants influence patients' experiences and satisfaction.[11] On the one hand, an increased understanding of such determinants enables professionals to tailor their care to the preferences and needs of different patient subgroups and to develop improvement programs addressing deficits in various settings. On the other hand, it might reveal the need to adjust for certain patient or clinic characteristics when interpreting patient evaluations, e.g., in the light of performance comparisons of health care providers.[12,13]

The objective of the present study was therefore twofold. First, we aimed to assess patients' experiences and satisfaction with fertility care, and second, we explored which clinic and patient characteristics are determinants of these two concepts.

Materials and methods

Setting

We conducted a cross-sectional study by postal questionnaire and two reminder rounds.[14] The reported results are part of a larger study which was approved by the Regional Review Board for Research on Human Subjects (CMO), Arnhem-Nijmegen (CMO no. 2004/193).

Study Population

We aimed to include a representative Dutch patient group, visiting 16 participating clinics for diagnostics or treatment of infertility. Of the 16 participating clinics, there was one university and one tertiary fertility clinic and six additional clinics offered IVF/ICSI treatment. These eight clinics were all large- or intermediate-sized teaching facilities and national health services funded, except for one smaller nonteaching private clinic. The remaining eight clinics were intermediate- or small-sized, and all but one nonteaching facilities. Every Dutch citizen has an either privately or publicly funded basic insurance coverage, which covered, during the study period, treatment

costs (but not medication costs) for intrauterine insemination (IUI) and ovulation induction (OI), as well as treatment and medication costs of three cycles of IVF/ICSI.
Potential participants were selected by means of each clinic's Diagnosis-Treatment Combination (DBC) registration database; in this nationwide financial registration method, infertile couples are identified with a specific F-code. An F-code could mean either an initial assessment of fertility, diagnostics, treatment, or counseling. Couples were apt for inclusion if they had an F-code any time in April, May, or June 2005. A computerized random sample of patients with an F-code was selected and invited to participate from each clinic's DBC database. The patient sample size was stratified according to clinic size (50, 150, or 550 patients) and encompassed ~75% of the clinic's fertility patients during the inclusion period. In total, 2,698 couples were invited. Patient couples were asked to complete the questionnaire concerning the study period (January 1, 2005–July 1, 2005). Couples were included if they completed both the questionnaire and the informed consent.

Questionnaire Development

Part 1: patients' experiences and satisfaction with fertility care

The patient questionnaire was constructed for a larger study on several aspects of clinical fertility care.[14] The present paper focuses on the questions concerning experiences and satisfaction with care, measured by a Dutch translation of Souter et al.'s ''questionnaire on patient satisfaction with the management of infertility''.[8] The original Souter questionnaire was systematically developed based on literature and focus group interviews with patients.

Part 2: determinants of patients' experiences and satisfaction with fertility care

The Souter questionnaire was complemented with questions regarding determinants of patients' experiences and satisfaction. Potential determinants were gathered from general and fertility literature[10,13,15–22] and expected to exist within three groups:

1. Patient's demographic characteristics (e.g., female age, couple's education level): Demographic characteristics were gathered from both partners by questions with open-end (e.g., ''age'') or categorical (e.g., ''household income'') response categories.
2. Patient's health status (e.g., level of anxiety, achieved pregnancy):
   The female partner was asked to fill out additional questions on anxiety and depression. Anxiety was measured by a 10-item short version of the State-Trait Anxiety Index (STAI)[23,24] and 12 additional fertility-related anxiety items, e.g., ''anxiety for treatment outcome,'' both on a 4-point scale (''almost never,'' ''sometimes,'' ''frequently,'' ''almost always''). Depression was measured by the Beck Depression Index for Primary Care (BDI-PC) using a 4-point scale of varying utterances[25]. ''Achieved pregnancy'' was defined as a positive pregnancy test in the study period. The questionnaire was piloted among 30 couples recruited by the Dutch Patients' Association for Infertility and adjusted where necessary.
3. Clinic's characteristics (e.g., clinic ID [1–16], presence of trained fertility nurses): A professional questionnaire regarding clinic characteristics and organization of fertility care[14] was sent to a coordinating gynecologist of each participating clinic.

Questions were either open ended (e.g., "number of fertility consultations/year") or categorical (e.g., "presence of trained fertility nurses").

Statistical Analysis

Dependent variables: patients' experiences and satisfaction with fertility care

Patients' experiences were considered to be dependent outcome variables. For reduction of dependent variables, exploratory factor analysis was performed on the Souter questionnaire. Five factors were discerned: 1) information provision; 2) emotional support; 3) waiting times; 4) doctor's attitude; and 5) organization of diagnostics. These five aspects are also mentioned in the Souter questionnaire, where patients are asked to rank them by their relative importance. A reliability analysis for internal consistency was performed for each factor (Cronbach $\alpha = 0.59–0.79$). Per patient, sum scores were calculated and dichotomized if distributions appeared skewed (i.e., categories containing <15% of all answers).

Independent variables: determinants of patients' experiences and satisfaction with fertility care

Confirmatory factor analysis and reliability assessment was performed for the questions of the STAI (Cronbach $\alpha = 0.91$) and BDI-PC (Cronbach $\alpha = 0.82$); both showed good internal consistency and "anxiety" and "depression" were carried forward as potential determinants. Exploratory factor analysis was performed on the additional anxiety questions, resulting in the factors "anxiety for treatment" (Cronbach $\alpha = 0.73$) and "anxiety for treatment outcome" (Cronbach $\alpha = 0.76$). Remaining anxiety items outside a factor were "anxiety for financial consequences" and "anxiety for relationship with partner." All independent variables were checked for colinearity. If a correlation coefficient of >0.6 was found, preference was given to the variable closest to actual outpatient performance. For example, a strong correlation (>0.8) was found between "clinic size" and "having special fertility consulting hours," and the latter was carried forward as the preferred variable.

All independent variables were then tested in an univariate analysis (cross tabulations, Wilcoxon rank sum test, independent-samples t test, and X^2-test) with the dependent variables concerning patient's experiences. The significantly associated (P<.100) variables that showed enough interclinic variation were included in a multilevel stepwise logistic regression analysis to explain differences in experiences. For this, a random coefficient model was composed using the levels "clinic" and "patient" in a Glimmix procedure in SAS (SAS forWindows v8.2). If no interclinic variation existed, regular multivariate regression analysis was performed. Significance for multivariate and multilevel regression analysis was set at P<.05.

Results

Response

In total, 1,499 couples (56%, range 47%–72%) completed the questionnaire and signed informed consent. Sociodemographic characteristics of the participants are shown in Table I (14). Mean female age was 32.8 years and male age 35.1 years. Only 1.9% (n=27) of couples were of non-Dutch origin, 53% had a high education level (>secondary school), and 94% had a more than modal household income (>1,760

euros/month gross); 73% suffered from primary infertility, and mean duration of infertility was 38 months.

Table I. Baseline characteristics of the participating couples (n=1499).

Characteristic	% of couples, or mean (SD)
Mean age	
Female	32.8 yrs (4.1 yrs)
Male	35.1 yrs (5.0 yrs)
Ethnic background[a]	
Dutch	98.1
Non-Dutch	1.9
Gross monthly family income (euros)[b]	
<1,100	1
1,100–1,760	4
1,760–2,750	22
>2,750	71.6
Education level per couple[c]	
Low	6
Intermediate	40
High	52.9
Type of infertility[d]	
Primary	72.8
Secondary	27.2
Mean duration of infertility[e]	38.4 mo (22.5 mo)

[a] Ethnic background of the couples was determined by the origin of both partners: Dutch = one or both partners are of Dutch origin; non-Dutch = both partners are not of Dutch origin.

[b] Gross monthly family income was categorized according to social security standards in 2005 and modal income in euros: <1,100 = less than Dutch modal income; 1,100–1,760 = Dutch modal income; 1,760–2,750 = up to 1,5 times Dutch modal income; > 2,750 = more than twice Dutch modal income.

[c] Education level of the couples was determined by the highest education level of both partners: low =primary or lower vocational education; intermediate = secondary or intermediate vocational education; high = higher professional education or university.

[d] Type of infertility was determined for the couple.

[e] Duration of infertility was defined as the period between the start of regular unprotected sexual intercourse and January 1 2005, the beginning of the study period.

Patients' experiences with fertility care

Table 2 shows patients' experiences with aspects of fertility care. In total, 38% of patients explicitly preferred a fertility clinic as opposed to a mixed clinic. A lack of continuity of care (i.e., being seen by more than one doctor) was indicated by 27% of patients. The aspects with the lowest amount of positive experiences were "emotional support," "waiting times," and "information provision." Emotional support by the team or an actual appointment with a social worker or psychologist was offered to 23% of patients, whereas 13% indicated they would actually want to talk to such professionals. Regarding consultations, 43% were held on scheduled time and mean waiting time was 18 minutes. However, 77% of patients seen late stated that this delay was acceptable. Written information was received by 70% of patients and 26% felt the amount was insufficient. Almost two-thirds (65%) of patients never received a clear plan for the future. Doctor's attitude was generally rated rather high. The majority of patients (>95%) felt that their doctor was listening, friendly, capable, sympathetic, and explaining. Approximately one in ten patients (11%) felt their doctor did not involve

them in decision making, one-fifth (19%) of patients felt that their doctor was not really interested in them as a person, and a similar number was not given enough opportunity to ask questions.

Ranking of aspects

The five aspects of care were ranked by patients for relative importance; ''Doctor's attitude'' was considered to be the most important aspect of care (36%), followed by ''information provision'' (29%), ''organisation of diagnostics'' (24%), ''waiting time'' (8%), and ''emotional support'' (4%).

Table II. Patients' experiences with fertility care, ranked per aspect.

	Response category ''yes''	
Aspect of care	**Number of patients**	**Percentage of patients**
Clinic's organization		
Preferred fertility clinic	573/1,459	38
Preferred mixed clinic	201/1,459	14
No preference	685/1,459	47
Experienced lack of continuity of care	348/1,300	27
Information provision		
Received written information	1,007/1,447	70
Received sufficient written information	354/1,362	26
Received sufficient explanation of possible causes	1,295/1,428	91
Received sufficient explanation of side-effects of medication	1,051/1,161	91
Received a clear plan for the future	452/1,276	35
Emotional support		
Was offered emotional support by the team	332/1,460	23
Was offered appointment with a social worker or psychologist	373/1,460	26
Wanted to talk to a social worker or psychologist at time of filling out questionnaire	182/1,379	13
Waiting times		
Seen on scheduled time	654/1,459	43
Delay acceptable	1,168/1,432	77
Doctor's attitude		
Did the doctor at your most recent visit:		
Behave politely	1,423/1,450	98
Appear good at his/her job	1,404/1,433	98
Listen to what you had to say	1,401/1,439	97
Explain things	1,379/1,445	95
Make it easy to ask questions	1,392/1,450	96
Show an interest in you as a person	1,114/1,379	81
Appear sympathetic	1,359/1,435	95
Let you take part in any decisions	1,154/1,299	89
Were there questions you would have liked to ask at the clinic but did not have the opportunity to do so	269/1,470	18
Organization of diagnostics		
Explained beforehand	1,364/1,410	90
Repeated excessively	84/1,394	6
Took too long to carry out	145/1,393	10
Took too long for results to come through	124/1,395	9

Determinants of Patients' experiences with care

Multilevel analysis showed significant determinants at four different domains (Table III). At patient level, three domains were repeatedly found: demographic characteristics (e.g., higher female age), type of received treatment (e.g., IVF), and health status. The latter concerned both mental health status (e.g., state anxiety) and general health status (e.g., achieved pregnancy). For example, achieving pregnancy was associated with positive experiences regarding emotional support (P=.014), doctor's attitude (P=.008), and organization of diagnostics (P=.003). At clinic level, only the domain ''organisation of care'' (e.g., presence of a separate fertility consultation hour) provided significant determinants. For example, having a separate fertility consultation hour was associated with positive experiences regarding emotional support (P=.019). Clinic size or ID was not associated with experiences.

Patient satisfaction and determinants of satisfaction with fertility care

Overall, 94% of patients were satisfied with current fertility services. Table IV shows the determinants associated with high satisfaction: suffering from secondary infertility (P=.048), less feelings of depression (P<.001), receiving IVF treatment (P=.016), achieving pregnancy (P<.001), and visiting a clinic with specialized fertility nurses (P=.019). Pregnancy outcome (e.g., ongoing pregnancy or not) was not significantly associated (P=.515) with satisfaction.

Discussion

This study assessed patients' experiences and satisfaction with several aspects of fertility care and explored the determinants influencing those experiences. Experiences were generally positive, which is similar to findings in a Scottish cohort of infertile patients.[8] Overall satisfaction with care was very high, which also corresponds with the literature.[8–10,26–28] Determinants of experiences and satisfaction could be found in both patient and clinic characteristics. Frequently found determinants were mental or general health status. Hall and Dornan[29] already described strong evidence that health status is a causal determinant of patient satisfaction. Also, higher levels of anxiety[17,18,30] and depression[16] are frequently found to be associated with dissatisfaction with care. For fertility care in particular, the role of social attributes, such as anxiety, about treatment as well as marital stress, has been reported to relate to negative patient experiences and lower satisfaction.[10,26,27] We found similar results: For each of the five aspects of care studied, lower scores of either anxiety or depression related to positive experiences with care. Moreover, lower depression scores were significantly associated with increased satisfaction.

The majority of studies on patient satisfaction with care unequivocally report that a good self-perceived health status or achieving a desired health outcome (in this case, pregnancy) is strongly associated with high satisfaction.[12,21,27,31,32] In accordance, we showed that achieving pregnancy was indeed associated with more positive experiences and higher overall satisfaction. Similarly, when the type of infertility was secondary (i.e., pregnancy was achieved before), it was associated with higher satisfaction ratings.

The only clinic characteristics found as determinants were the ''presence of trained nurses'' and ''specialized fertility consulting hours''. This reflects the fact that specific

patient-centered attention and tailored care may be accountable for positive experiences and high satisfaction.[11,15,33]

We assume this is also the reason why specific types of received treatment (i.e., IUI and IVF/ICSI treatment) were repeatedly found as a determinant. It might appear strange that the most intensive treatment types (as opposed to, e.g., expectant policy) were found to be determinants. However, invasive and time-consuming types of treatment are provided within specifically designed outpatient care, involving more extensive information provision and contact hours with highly specialized personnel. Patients' preferences for more continuity and personal care are thus probably met.[11,34,35] This is also described by others, who found that emotional support and a good relationship between staff and patient was a major determinant of patient satisfaction.[15,28,33] Considering the high drop-out rates (i.e., 17%–65%) known to exist in IVF/ICSI treatment, it would probably prove worthwhile for clinics to invest in better emotional support to enable couples to go through treatment cycles confidently.[36–42]

Global differences in reimbursement of fertility care could hamper adoption of such beneficial organizational elements. However, because clinic characteristics might be the only determinants that can be easily influenced by care providers (as opposed to, e.g., type of fertility), efforts should be made to adapt these, independent of reimbursement systems. Even with limited financial means, small adjustments in existing clinic structure could be easily achieved by redesigning current care structure. For example, incidental or randomly planned fertility consultations in mixed obstetrics/gynecology outpatient clinics could be streamlined by clustering fertility patients into separate consultation hours where nursing staff remains of stable composition. This nursing staff is preferably extra educated or subspecialized within infertility care or, if reimbursement is insufficient, staff can be enabled to get educated "on the floor."

Experiences Versus Satisfaction

The first initiatives at incorporating patients' opinions in health care assessments concerned mainly satisfaction ratings. However, these carry the disadvantage of being subject to emotions, thus presenting only limited pictures. This generally results in high satisfaction ratings, even when patients simultaneously indicate problems with several aspects of care.[15,43–45] Measuring patients' concrete experiences with specific aspects of care (e.g., "did your doctor give you the opportunity to ask questions?") rather than satisfaction (i.e., "were you satisfied with your doctor?") therefore provides a more reliable indication of the quality of care delivered, because it provides insight into the processes in need of improvement.[15,44,46] The same applies to assessing background factors of patients' satisfaction with care, as opposed to patients' experiences with care.[44] The concepts "experience" and "satisfaction," however, remain to be intertwined.[47,48] Ware et al.[49] made an early attempt at theoretic work by distinguishing satisfaction "reports" (i.e., experiences) from satisfaction "ratings"[50], and only gradually the term "experiences" was introduced to discriminate between the two. Still, "satisfaction" is used alternately for the same concept; for example, the Souter questionnaire is reported by the original authors to assess "patient satisfaction," although, in our opinion, it actually assesses mostly experiences while also including a single question on satisfaction with care. The high ratings of

satisfaction and experience described here are in correspondence with other fertility literature[8–10,26–28], However, most of these studies are performed in countries which also have fairly good ART reimbursement systems. Patients' expectations of fertility care might therefore be lower compared with countries where patients have to pay directly for treatment, because expectations are assumed to influence the way health care delivery is experienced.[51]

Limitations

There are also some limitations to this study. First, the participants in the study were relatively highly educated and belonged to high-income households of primarily Dutch origin. This rather homogeneous background could have caused the absence of determinants such as age, ethnicity, education, and income, which are known from literature to relate to satisfaction.[12,30,52,53] The use of an extensive written questionnaire could have induced a sample bias, because non-Dutch or lower-class couples are less apt to participate owing to inadequate language proficiency. We were unable to perform a nonresponder analysis, because of absent demographic information on the nonresponder group, although this could have provided information on the extent of sample bias. However, the demographic background of the study group was similar to those in a previous survey study in the same region, which had a higher response rate.[44]

Second, earlier studies found gender differences in experiences and satisfaction assessments.[27,54,55] Most of our questionnaire items were exclusively aimed at the female partner. Future research could benefit from using separate questionnaires for both partners.[14]

Finally, the major disadvantage of the study is the cross-sectional study design, which can only demonstrate associations, not causality. To demonstrate any causality, future studies should be longitudinal in design.

Conclusion

In conclusion, the present study provides us with an increased understanding of the determinants of patients' experiences and satisfaction with comprehensive fertility care. Determinants could be found at both patient and clinic levels. Couples clearly seem to prefer continuity of care, including the opportunity to connect with a small group of health care providers, versus the ''clinic'' type of care. These results enable fertility professionals to tailor their care to specific subgroups of patients and adjust their organization of care where needed. Moreover, this study underscores the need to investigate whether case-mix correction is necessary when interpreting patient surveys on experiences and satisfaction with care, because patient and clinic characteristics can strongly influence the way health care delivery is experienced.

Table III. Determinants of patients' positive experiences with several aspects of fertility care—results of a multilevel logistic analysis.

Aspect of fertility care	Determinant		OR [95% CI]	P value	Determinant associated with positive experiences
Information provision	Mental health status	(Depression)	0.88 [0.84–0.92]	< .001	Less depressive feelings
	Type of treatment	(IUI)	1.88 [1.33–2.64]	< .001	Received IUI treatment
		(IVF/ICSI)	1.82 [1.25–2.65]	.002	Received IVF/ICSI treatment
Emotional support	Health status	(Pregnancy)	1.48 [1.08–2.01]	.014	Achieved pregnancy
	Mental health status	(Anxiety)	0.93 [0.90–0.96]	< .001	Less anxious feelings (state anxiety)
		(Depression)	0.90 [0.85–0.96]	.002	Less depressive feelings
	Type of treatment	(IVF/ICSI)	2.53 [1.67–3.85]	< .001	Received IVF/ICSI treatment
	Organization of care	(Consultation hours)	2.12 [1.13–3.96]	.019	Presence of specialized fertility consultation hours
Waiting times	Mental health status	(Depression)	0.92 [0.88–0.97]	< .001	Less depressive feelings
	Type of treatment	(IUI)	1.43 [1.04–1.96]	.028	Received IUI treatment
		(IVF/ICSI)	1.70 [1.12–2.58]	.013	Received IVF/ICSI treatment
Doctor's attitude	Health status	(Pregnancy)	1.51 [1.12–2.04]	.008	Achieved pregnancy
	Mental health status	(Anxiety)	0.97 [0.95–0.99]	.024	Less anxious feelings (state anxiety)
	Type of treatment	(IVF/ICSI)	1.53 [1.05–2.24]	.028	Received IVF/ICSI treatment
Organization of diagnostics	Demographics	(Female age)	1.57 [1.03–2.38]	.010	Younger female age
	Health status	(Pregnancy)	1.85 [1.19–2.94]	.003	Achieved pregnancy
	Mental health status	(Anxiety)	1.79 [1.22–2.56]	< .001	Less anxious feelings(financial consequences)
	Type of treatment	(IVF/ICSI)	1.03 [1.01–1.11]	.007	Received IVF/ICSI treatment

Table IV. Determinants of patients' satisfaction with fertility care—results of a multilevel logistic analysis.

	Determinant		OR [95% CI]	P value	Determinant associated with higher satisfaction
Satisfaction	Demographics	(Type of subfertility)	1.97 [1.01–3.87]	.048	Secondary subfertility
	Health status	(Pregnancy)	3.70 [1.85–7.38]	<.001	Achieved pregnancy
	Mental health status	(Depression)	0.88 [0.82–0.95]	<.001	Less depressive feelings
	Type of treatment	(IVF/ICSI)	2.79 [1.21–6.39]	.016	Received IVF/ICSI treatment
	Organization of care	(Trained nurses)	2.08 [1.13–3.85]	.019	Presence of trained fertility nurses

CI = confidence interval; IUI = intrauterine insemination; IVF/ICSI = in vitro fertilization/intracytoplasmatic sperm injection; OR = odds ratio.

References

1. Jenkinson C, Coulter A, Bruster S. The Picker Patient Experience questionnaire: development and validation using data from in-patient surveys in five countries. Int J Qual Health Care 2002;14:353–8.
2. Jenkinson C. Assessment and evaluation of health and medical care. Buckingham, U.K: Open University Press, 1997.
3. Cleary PD, Edgman-Levitan S, Walker JD, Gerteis M, Delbanco TL. Using patient reports to improve medical care: a preliminary report from 10 hospitals. Qual Manag Health Care 1993;2:31–8.
4. Cleary PD, Edgman-Levitan S. Health care quality. Incorporating consumer perspectives. JAMA 1997;278:1608–12.
5. Gerteis M, Edgman-Levitan S, Daley J, Delbanco T. Through the patient's eyes. San Francisco, CA: Jossey-Bass, 1993.
6. Wensing M, Elwyn G. Research on patients' views in the evaluation and improvement of quality of care. Qual Saf Health Care 2002;11:153–7.
7. Cleary PD. The increasing importance of patient surveys. Now that sound methods exist, patient surveys can facilitate improvement. BMJ 1999;319:720–1.
8. Souter VL, Penney G, Hopton JL, Templeton AA. Patient satisfaction with the management of infertility. Hum Reprod 1998;13:1831–6.
9. Schmidt L. Infertile couples' assessment of infertility treatment. Acta Obstet Gynecol Scand 1998;77: 649–53.
10. Malin M, Hemmink E, Raikkonen O, Sihvo S, Perala ML. What do women want? Women's experiences of infertility treatment. Soc Sci Med 2001;53: 123–33.
11. Cleary PD, McNeil BJ. Patient satisfaction as an indicator of quality care. Inquiry 1988;25:25–36.
12. Hargraves JL, Wilson IB, Zaslavsky A, James C, Walker JD, Rogers G, et al. Adjusting for patient characteristics when analyzing reports from patients about hospital care. Med Care 2001;39: 635–41.
13. Young GJ, Meterko M, Desai KR. Patient satisfaction with hospital care: effects of demographic and institutional characteristics. Med Care 2000;38:325–34.
14. Mourad SM, Hermens RP, Cox-Witbraad T, Grol RP, Nelen WL, Kremer JA. Information provision in fertility care: a call for improvement. Hum Reprod 2009.
15. Jenkinson C, Coulter A, Bruster S, Richards N, Chandola T. Patients' experiences and satisfaction with health care: results of a questionnaire study of specific aspects of care. Qual Saf Health Care 2002;11:335–9.
16. Lee DS, Tu JV, Chong A, Alter DA. Patient satisfaction and its relationship with quality and outcomes of care after acute myocardial infarction. Circulation 2008;118:1938–45.
17. Britton JR. Global satisfaction with perinatal hospital care: stability and relationship to anxiety, depression, and stressful medical events. Am J Med Qual 2006;21:200–5.
18. Herrera-Espineira C, Rodriguez Del Aguila MM, Rodriguez Del CM, Valdivia AF, Sanchez IR. Relationship between anxiety level of patients and their satisfaction with different aspects of healthcare. Health Policy 2009;89:37–45.
19. Jung HP, Baerveldt C, Olesen F, Grol R, Wensing M. Patient characteristics as predictors of primary health care preferences: a systematic literature analysis. Health Expect 2003;6:160–81.
20. Jones R, Pearson J, McGregor S, Gilmour WH, Atkinson JM, Barrett A, et al. Cross sectional survey of patients' satisfaction with information about cancer. BMJ 1999;319:1247–8.
21. Westaway MS, Rheeder P, Van Zyl DG, Seager JR. Interpersonal and organizational dimensions of patient satisfaction: the moderating effects of health status. Int J Qual Health Care 2003;15: 337–44.
22. Schmidt L, Holstein BE, Boivin J, Sangren H, Tjornhoj-Thomsen T, Blaabjerg J, et al. Patients' attitudes to medical and psychosocial aspects of care in fertility clinics: findings from the Copenhagen Multi-centre Psychosocial Infertility (COMPI) research programme. Hum Reprod 2003;18:628–37.
23. Spielberger R, Gorsuch R, Lushene R. STAI manual for the State-Trait Anxiety Inventory 1970. Palo Alto, CA: Consulting Psychologists, 1970.

24. Van Der Ploeg H, Defares P, Spielberger C. Handleiding bij de Zelf-Beoordelings Vragenlijst ZBV: Een nederlandstalige bewerking van de Spielberger State-Trait Anxiety Inventory, STAI-DY. Lisse. The Netherlands: Swets and Zeitlinger, 1980.
25. Beck AT, Guth D, Steer RA, Ball R. Screening for major depression disorders in medical inpatients with the Beck Depression Inventory for Primary Care. Behav Res Ther 1997;35:785–91.
26. Sabourin S, Wright J, Duchesne C, Belisle S. Are consumers of modern fertility treatments satisfied? Fertil Steril 1991;56:1084–90.
27. Schmidt L, Holstein BE, Boivin J, Tjornhoj-Thomsen T, Blaabjerg J, Hald F, et al. High ratings of satisfaction with fertility treatment are common: findings from the Copenhagen Multi-centre Psychosocial Infertility (COMPI) research programme. Hum Reprod 2003;18:2638–46.
28. Halman LJ, Abbey A, Andrews FM. Why are couples satisfied with infertility treatment? Fertil Steril 1993;59:1046–54.
29. Hall JA, Dornan MC. Patient sociodemographic characteristics as predictors of satisfaction with medical care: a meta-analysis. Soc Sci Med 1990;30:811–8.
30. Rahmqvist M. Patient satisfaction in relation to age, health status and other background factors: a model for comparisons of care units. Int J Qual Health Care 2001;13:385–90.
31. Sundby J, Olsen A, Schei B. Quality of care for infertility patients. An evaluation of a plan for a hospital investigation. Scand J Soc Med 1994;22:139–44.
32. Nguyen Thi PL, Briancon S, Empereur F, Guillemin F. Factors determining inpatient satisfaction with care. Soc Sci Med 2002;54:493–504.
33. Caljouw MA. van BM, Boer F. Patient's satisfaction with perioperative care: development, validation, and application of a questionnaire. Br J Anaesth 2008;100:637–44.
34. Saultz JW, Albedaiwi W. Interpersonal continuity of care and patient satisfaction: a critical review. Ann Fam Med 2004;2:445–51.
35. Sans-Corrales M, Pujol-Ribera E, Gene-Badia J, Pasarin-Rua MI, Iglesias-Perez B, Casajuana-Brunet J.

Chapter 7

A multi-faceted strategy to improve the use of national fertility guidelines; a cluster-randomized controlled trial

S.M. Mourad
R.P.M.G. Hermens
J. Liefers
R.P. Akkermans
G.A. Zielhuis
E.M. Adang
R.P.T.M. Grol
W.L.D.M. Nelen
J.A.M. Kremer

Submitted.

Abstract

Introduction: Proper use of clinical practice guidelines can decrease variation in care between settings. However, actual use of fertility guidelines is suboptimal and in need of improvement. Hence, a cluster-randomized controlled trial was designed to study the effects of two strategies to implement national Dutch guidelines on comprehensive fertility care.

Materials and methods: Design: Sixteen fertility clinics participated in the trial. A minimal, professional oriented implementation strategy of audit and feedback was tested versus a maximal multi-faceted strategy that was both professional as well as patient-oriented.
Outcome measures: The extent of adherence to guideline-recommendations, reflected in quality indicator scores. To get insight into unwanted side-effects, patient anxiety and depression scores were gathered as secondary outcomes. Data collection encompassed medical record search, patient and professional questionnaires.

Results: 1499 couples were included at baseline and 1396 at after-measurement. No overall significant improvement in indicator scores was found for either strategy. Secondary outcomes did not differ for both groups, selected anxiety scores were lower in the maximal intervention group. Process evaluation of the trial revealed positive patient experiences with the intervention material. Professionals' appreciation of intervention elements varied, and execution of the multi-faceted strategy appeared incomplete.

Discussion: Absence of an intervention effect may be due to the nature of the strategies, incomplete execution or flaws in study design. Process evaluation data raises the question whether professionals should be the only actors responsible for guideline implementation. This study therefore contributes to an increased understanding of fertility guideline implementation in general, and the role of patients in particular.

Introduction

The burden of infertility weighs heavily on approximately 80 million couples worldwide.[1,2] For western Europe, this affects around 10-15% of couples of reproductive age[3-6], of which an estimated 50% seeks medical assistance.[4] Providing those infertile couples with best available health care, i.e. consensus and evidence-based diagnostic tests and treatment options, is the common aim of several national and international societies of fertility professionals. For this purpose, they developed clinical practice guidelines to guide health care providers during daily practice, while at the same time decreasing variation in care between settings. However, the mere existence of such guidelines does not automatically imply that they are widely spread nor commonly used. Dissemination of new guidelines should therefore ideally be followed by robust implementation efforts. However, research on the implementation of fertility guidelines in particular is scarce. This is remarkable, as the diagnostics and treatment commonly used for infertility, lead to an extensive use of healthcare resources and are associated with substantial physical and psychological consequences for the patients involved.[7-11]

It is commonly known in guideline implementation research that there is no 'magic bullet' for successful implementation of every clinical problem or in every practice setting. For example, literature is still inconclusive regarding the effects of multifaceted versus single interventions for guideline implementation.[12-16] The most frequently studied interventions encompass audit and feedback on current practice, dissemination of educational materials, reminders and the organization of educational meetings or outreach visits, which all seem to have only small to moderate effects on the improvement of professional performance.[17-19] However, mainly individual professionals are targeted in these interventions. The role of patients in guideline implementation is, surprisingly, still rather unexplored and evidence is therefore scarce.[20-22] As fertility patients are generally young, critical towards their care providers and thus sometimes regarded by professionals as one of many barriers to guideline implementation[23], we hypothesize patients could be just the allies we need for successful implementation of fertility guidelines.

Hence, we designed a cluster-randomized controlled trial (c-RCT) to study the effects of two different strategies to improve the use of recommendations in national Dutch fertility guidelines on comprehensive clinical fertility care. Literature was reviewed to identify barriers for the implementation of clinical practice guidelines in reproductive medicine and fertility care in particular[23-31], as a prospective identification of barriers is still assumed to lead to better adaptation of interventions.[32] Based on this literature, barriers were expected to be found in the contexts of the guideline itself, organizational aspects of care, patient characteristics and professional characteristics. Our strategies were subsequently tailored to these barriers. We hypothesized that a maximal implementation strategy consisting of a multi-faceted intervention, tailored to barriers from literature and directed at both professionals as well as patient couples, would prove to be more effective than a minimal implementation strategy consisting of a single, professional-oriented intervention of audit and feedback. We assume these effects can be seen at cluster level. To increase our understanding of

factors influencing the impact of the implementation strategies, a process evaluation of the trial was also performed.

Material and methods

Study design

We performed a cluster-randomized controlled trial in 16 fertility clinics from the Fertility Network East, a knowledge and clinical experiences exchanging network of fertility clinics in the eastern region of the Netherlands. The study was approved for all clinics by 'the Regional Review Board for Research with Human Subjects (CMO) Region Arnhem-Nijmegen' (CMO no. 2004/193), as well as two local research ethics committees. The trial was registered with ClinicalTrials.gov (ID NCT00119925) and results are reported according to the CONSORT–statement for cluster-randomized trials.[33]

Randomization

Fertility clinics instead of couples were the unit of randomization to avoid cross-over contamination of both intervention strategies; clinics were assigned to either a minimal or a maximal implementation strategy. Prior to randomization to one of the implementation strategies, participating clinics were stratified according to clinic size (small, medium large) and treatment facilities (with or without IVF/ICSI facilities) (see figure 1). Two independent research associates performed the randomization procedure by drawing blinded envelopes. Allocation sequence was concealed until the interventions were assigned. Usual care measured at baseline served as a control for both groups. Figure 1 shows a flow-chart of clinics and included patients.

Blinding

All patients and professionals were blinded to group assignment and remained unaware of intervention contents. The research group, but not the trained data extraction personnel was aware of group assignment.

Participants

Clinics

In total, sixteen clinics of the Fertility East Network were invited and agreed to participate. There was one academic and one tertiary care clinic, both running a licensed IVF laboratory, of which there are only 13 in total in the Netherlands. Seven of the sixteen clinics offered secondary care, as an autonomous, satellite (i.e. performing an IVF/ICSI cycle up to the ovarian stimulation phase) or transport IVF/ICSI clinic satellite (i.e. performing an IVF/ICSI cycle up to the oocyte pick-up phase). These clinics are also teaching clinics and of large or intermediate size; the other seven clinics are smaller, non-teaching facilities. In total, 15 clinics are in the national health system, and one of the smaller secondary care clinics is a private clinic. The different types of clinics were chosen to reflect average national fertility care.

Patients

To include a representative patient group, potential participating couples were retrospectively selected in each clinic by means of the clinics' financial registration database (Diagnosis Treatment Combination code). In this nation-wide registration

system, patients undergoing diagnostics or treatment for infertility are identified with a specific Fertility-code (F-code). A baseline measurement was performed which included a random sample of 1499 couples that had an active F-code anytime between the 1st of April 2005 and July 1st 2005 and concerned the fertility care these couples received in the period between the 1st of January 2005 and July 1st 2005. For the after-measurement, couples were apt for inclusion if they had an active F-code anytime between the 1st of December 2007 and March 1st 2008. The after-measurement focused on the care these couples received in the period between the 1st of September 2007 and March 1st 2008. From each clinic, a random sample of eligible couples (n=50, 150 or 500, stratified according to clinic size) was invited to participate in data-collection for the study (postal survey). Patients in each study arm were included according to the 'intention to treat' principle, i.e. independent from diagnosis or treatment type. This procedure was followed for both the baseline and after-measurement. Eligible couples were sent a questionnaire which included an informed consent form concerning the use of both medical data from patient records as well as from the questionnaire.

Study Interventions

Minimal strategy; Audit and Feedback

The minimal implementation strategy was entirely directed at professionals, and in June 2007 introduced in each of the eight clinics in the minimal intervention arm. This strategy consisted of professional audit and feedback. Audit encompassed results of the baseline assessment of the clinics' scores to previously developed quality indicators[34], regarding care provided in the period January 2005-July 2005. Based on these indicator scores, clinic-specific feedback reports regarding current care were formulated by the study group. Each of these minimal strategy clinics was sent a sufficient number of these feedback reports for dissemination among its fertility professionals (i.e. gynaecologists, fertility doctors, nurses, physician assistants). An instruction letter for interpretation was enclosed in each report. Per quality indicator, feedback on current care was given by means of a bar-chart showing the total range of performance on a scale from 0%-100% including the median adherence of all participating clinics. Feedback was kept anonymously and each individual clinic received a feedback report with a clear marking in the bar chart of their own performance. After sending these feedback reports, no further contact was established with the minimal intervention clinics until the data collection for the after-measurement was started up.

Maximal strategy; multi-faceted intervention

The maximal implementation strategy was multifaceted of nature, and tailored to barriers known from literature as well as to baseline performance; the strategy was directed at both professionals and patient couples.

The strategy included the following four professional oriented elements:

i) Audit and feedback discussions

Similar to the minimal strategy, clinic-specific feedback reports on current care were developed and sent to the eight maximal intervention clinics in May 2007. After circa

three weeks, a multidisciplinary (e.g. addressing gynecologists, fertility doctors, residents, fertility nurses, quality officers) one hour meeting was organized in each clinic, where the feedback report was presented and commented on by one of the authors (S.M.). During each meeting, the clinic's performance was discussed in the light of the other 15 clinics' performances, and possibilities for quality improvement were highlighted.

Moreover, professionals in each of the maximal strategy clinics were provided with the following implementation tools:

i) List of suggested tools for implementation of the guidelines at local level

Per indicator, a specific tool or action was provided to improve performance. For example, the proposed action to improve 'history-taking of the couple' consisted of 'the development and introduction in the outpatient record of a systematic history-taking form'.

Besides these professional oriented interventions, a patient oriented intervention was introduced in these maximal strategy clinics:

ii) Leaflet on Shared Decision Making (SDM) in the fertility consultation

The leaflet on SDM explained the general principles of SDM and contained a suggested literature list. It also provided practice examples for fertility care in which patient preferences do not always match guideline contents, and how SDM could be helpful to reach bilateral agreement in policy (e.g. concerning single versus double embryo transfer or the need for an expectant period before start of assisted reproduction).

iii) Patient information checklists

Checklists for the provision of patient-information were disseminated for professionals to use in the consultation room. These checklists consisted of laminated sheets that contained patient information items to be discussed with patients and recommended in the guidelines (e.g. the chances of pregnancy after transfer of one versus two embryos, the side-effects of medication for ovulation induction).

iv) Educational patient leaflets

Based on the fertility guidelines and accompanying quality indicators, three educational patient leaflets were developed concerning; 1) the initial assessment of fertility, male infertility, endometriosis, tubal pathology and idiopathic infertility; 2) ovulation induction, intra uterine insemination and ovarian hyper stimulation syndrome; 3) IVF/ICSI treatment and ovarian hyper stimulation syndrome. These leaflets explained the background and contents of the professional guidelines in lay language. The text moreover aimed to encourage patients to start a dialogue with their doctor regarding diagnostic testing, treatment plan and treatment itself, while promoting the concept of shared decision making. The leaflets contained a prompt sheet for questions during consultation and also referred to the national fertility guidelines published on the website of the Dutch Society for Obstetrics and Gynaecology. Professionals were asked to distribute the relevant leaflets among the infertility patients consecutively visiting their (outpatient) clinic in the implementation period (June 2007-January 2008).

Outcomes

The effectiveness of both implementation strategies was reflected in the scores on a pre-developed set of guideline based quality indicators[34]. The studied guidelines

included nine fertility guidelines of the Dutch Society for Obstetrics and Gynaecology (NVOG): Initial assessment of fertility, anovulation, male infertility, tubal pathology, endometriosis, premature ovarian failure (POF), intra uterine insemination (IUI), indications for in vitro fertilization (IVF) treatment, ovarian hyperstimulation syndrome (OHSS). Moreover the model protocol of the Dutch embryo Act was considered, which includes clinical statements on the provision of IVF treatment. The indicators were systematically developed structure and process indicators for comprehensive clinical fertility care. They cover topics like 'indications for treatment', 'diagnostic procedures', 'treatment procedures' and 'patient information'. Indicators are scored dichotomously, in which the value '1' reflects 'adherence' and value '0' reflects 'non-adherence' to a defined guideline recommendation. All indicators were tested during the baseline measurement for several quality criteria (i.e. measurability, reliability, applicability, improvement potential, discriminatory capacity, complexity and case-mix stability, thus exploring their value as instruments for monitoring and improving clinical performance.[34] After the baseline measurement indicators with high applicability were considered adequate primary outcome measures to reflect the degree of adherence to guideline recommendations for the c-RCT.[35] Indicators which also have high improvement potential are suitable to assess changes in adherence, whereas indicators with low improvement potential are only suitable to monitor adherence in time. To get insight into any unwanted side-effects of the applied strategies, we included at patient level 'anxiety' and 'depression' as secondary outcome measures.

Data collection

Data collection to calculate quality indicator scores was performed from either medical records (e.g. process indicators concerning treatment aspects), a professional questionnaire (e.g. structure indicators concerning clinic's structure) or a patient questionnaire (e.g. process indicators concerning patient information provision).[36] Medical record extraction was performed by three trained data collectors who entered data in digital forms, specifically designed to enhance systematic and complete data collection by using computerized algorithms for data entry. Data collectors followed an intensive one-month training and reliably assessed a series of 30 test records before starting official data collection. During data-collection, two independent reviewers abstracted a random sample of 10% (n=32) of medical records from 2 participating clinics. The extent of agreement between these data reviewers on the level of process indicator scores, corrected for chance, was calculated using Cohen's kappa coefficient[37]. Reliability among the data collectors was substantial, reflected in average Cohen's Kappa-coefficients of 0.86 (baseline) and 0.82 (after-measurement) (range 0.48 -1.0).

Data for the secondary outcome measures were gathered by the patient questionnaire. Anxiety was measured by a 10 item short version of the State Trait Anxiety Index (STAI)[37,38] and n=12 additional infertility-related anxiety items, e.g. 'anxiety for treatment outcome, both on a 4-point scale ('almost never', 'sometimes', 'frequently', 'almost always'). Depression was measured on a 4 point scale by the Beck Depression Index for Primary Care (BDI-PC)[40].

Process evaluation of intervention elements

We applied the process-evaluation framework described by Hulscher et al.[41] to evaluate both exposure to as well as each stakeholder's experience with the several elements of the interventions. This was done by means of a professional questionnaire and an addendum to the patient questionnaire in the after-measurement.

Sample size

The study was designed to provide at least 80% power in order to detect a difference of 15% in guideline adherence between the two study arms (70% for the minimal strategy and 85% for the maximal strategy) at the 0.05 two-sided significance level. Considering an intracluster-correlation coefficient (ICC) of 0.15 and assuming an average 30 infertile couples seen per professional and an average 2 á 3 professionals per clinic (43 professionals in total), 1290 couples (43*30) were needed for both the baseline measurement as well as the after-measurement.

Analysis

To analyze the effectiveness of the two implementation strategies, we assessed the proportion of patients that was treated in accordance with the guidelines, which is reflected in quality indicators scores before and after the implementation period in both intervention arms. The difference in indicator scores was analyzed with adjustment for clustering of patients within clinics. Therefore, for each indicator, multi-level logistic regression analyses were performed in which 'intervention arm' acted as the independent variable and 'indicator score' as the dependent variable. Analyses were based on the modified intent to treat principle, meaning all participants were included in the arm to which they were originally assigned, regardless of whether they completed the intervention given to the arm. Chi^2 statistic and Crosstabs were used to determine differences in scores within either baseline or the after-measurement. If significant differences (set at $p<0.05$) existed at baseline level for either specific indicator scores or demographic characteristics, these differences were corrected for by taking baseline scores up as a covariate in the final multi-level model composed to assess differences between the two intervention arms. The statistical programme SAS for Windows V8.2 was used to compose a multi-level model correcting for clustering in the different clinics.

Results

Participant flow

At baseline, 726 couples were included for the maximal implementation strategy and 773 for the minimal implementation strategy, for the after-measurement these numbers were 697 and 696 respectively (see figure 1).

Comparability of data at baseline

Demographic characteristics of both intervention arms at both measurement periods are shown in table I. For most characteristics, the arms were comparable.

There were significant differences between the minimal and maximal intervention arms at baseline, for the variables 'type of infertility' (68% versus 77% primary infertility; p=0.000) and 'duration of infertility' (35 versus 41 months; p=0.000).

Figure 1

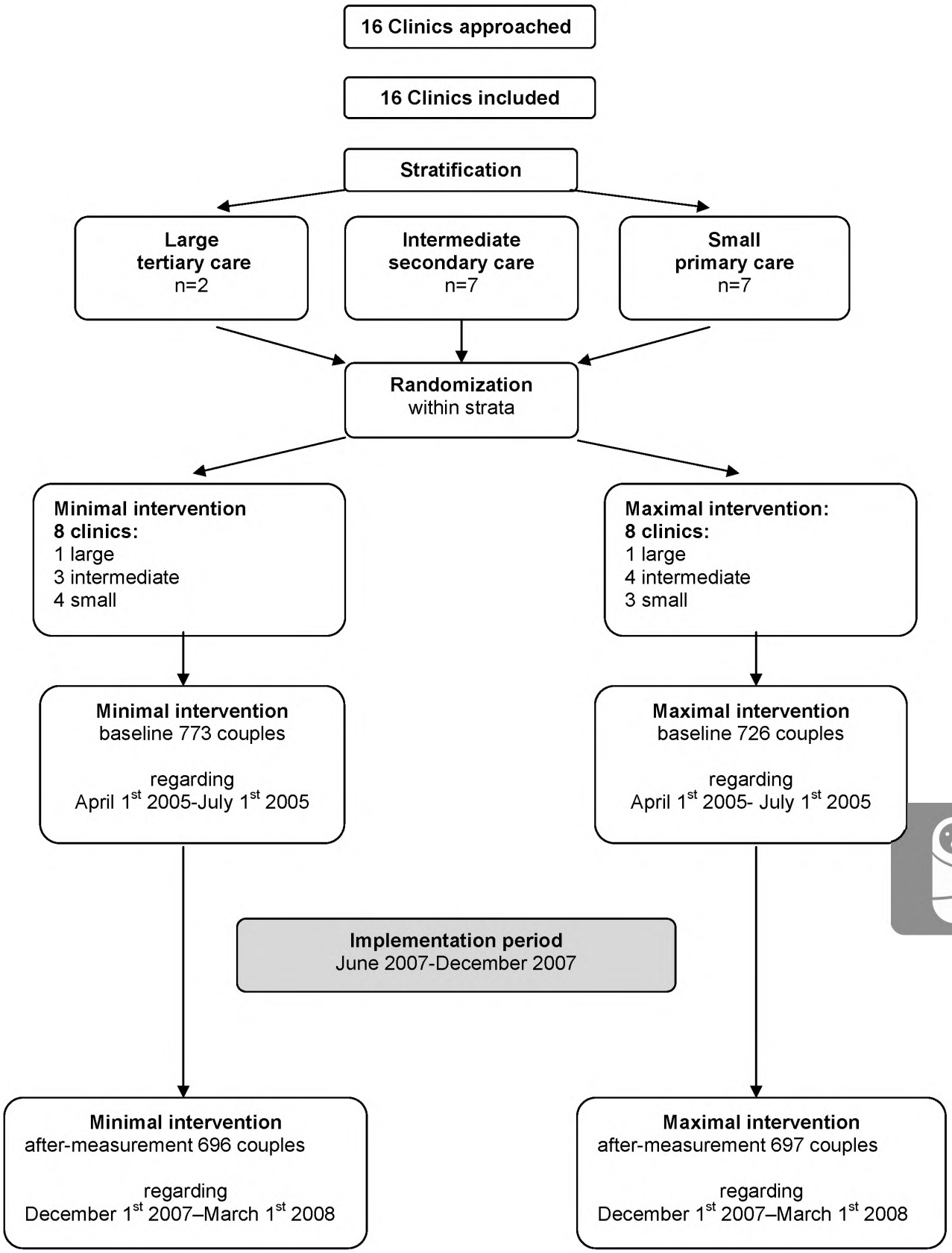

16 Clinics approached
16 Clinics included
Stratification
Large
tertiary care
n=2
Intermediate
secondary care
n=7
Small
primary care
n=7
Randomization
within strata
Minimal intervention
8 clinics:
1 large
3 intermediate
4 small
Maximal intervention:
8 clinics:
1 large
4 intermediate
3 small
Minimal intervention
baseline 773 couples
regarding
April 1st 2005-July 1st 2005
Maximal intervention
baseline 726 couples
regarding
April 1st 2005- July 1st 2005
Implementation period
June 2007-December 2007
Minimal intervention
after-measurement 696 couples
regarding
December 1st 2007–March 1st 2008
Maximal intervention
after-measurement 697 couples
regarding
December 1st 2007–March 1st 2008

Table I demographic characteristics of the participating couples

Characteristics	Baseline measurement (n=1499)		After-measurement (n=1393)	
	Minimal arm (n=726)	Maximal arm (n=773)	Minimal arm (n=696)	Maximal arm (n=697)
Mean age in years (SD)				
Female	32.48 (SD 4.1)	32.85 (SD 4.2)	33.06 (SD 4.5)	33.37 (SD 4.7)
Male	35.11 (SD 4.9)	35.56 (SD 5.1)	35.53 (SD 5.1)	36.31* (SD 6.1)
Ethnic background (%)[a]				
Dutch	98.4	97.8	98.5	98.4
Non-Dutch	1.6	2.2	1.5	1.6
Gross monthly family income in € (%) [b]				
<1100	1.3	1.9	0.8	1.5
1100-1760	4.4	4.8	2.3	4.3
1760-2750	20.9	23.4	14.8	15.5
>2750	73.3	69.9	82.1	78.6
Education level per couple (%)[c]				
Low	7.1	6.1	3.1	5.7
Intermediate	41.1	40.2	36.6	36.4
High	51.9	53.6	60.3	57.9
Type of infertility (%)[d]				
Primary	68.4	77.0*	67.8	69.5
Secondary	31.6	23.0*	32.2	30.5
Mean duration of infertility in months (SD)[e]	34.7 (SD 20.4)	41.4* (SD 24.2)	35.8 (SD 23.7)	38.7* (SD 26.1)

* significant difference between both arms at the p=0.05 level.

[a] Ethnic background of the couples was determined by the origin of both partners: Dutch = one or both partners are of Dutch origin; non-Dutch = both partners are not of Dutch origin.

[g] Gross monthly family income was categorized according to social security standards in 2005 and modal income in euros:<1100 less than Dutch modal income; 1100-1760 Dutch modal income; 1760-2750 up to 1,5 times Dutch modal income; >2750 more than twice Dutch modal income.

[c] Education level of the couples was determined by the highest education level of both partners: low = primary or lower vocational education; intermediate = secondary or intermediate vocational education; high = higher professional education or university.

[d] Type of infertility was determined for the couple.

[e] Duration of infertility was defined as the period between the start of regular unprotected sexual intercourse and the beginning of the study period: respectively January 1st 2005 (for the baseline measurement) or September 1st 2007(for the after-measurement).

Outcomes and estimation

PRIMARY OUTCOMES

Table II shows the indicator scores for both the baseline as the after-measurement. The indicators with low improvement potential at baseline showed in the after-measurement the same high adherence percentages, meaning care remained of high standards. This is for example true for the indicator 'Couple's history-taking should cover at least: age of both partners, duration of infertility, type of couple's infertility'.
The remaining indicators with high improvement potential at baseline, showed mixed effects in the after-measurement (table II).
Only for two indicators, concerning 'the aim of the initial assessment of fertility to result in a diagnosis and prognosis' and 'lifestyle advice concerning smoking', there was a significant surplus value of the maximal intervention compared to the minimal intervention arm. For one indicator, concerning the monitoring of ovarian response by transvaginal ultrasound in case of IUI in the stimulated cycle, there was a significant less decrease in adherence to the guideline when comparing minimal and maximal intervention arms.

SECONDARY OUTCOMES

Both at baseline as well as in the after-measurement, there were no significant differences in BDI-PC and STAI scores for the minimal and maximal intervention arms (data not shown). Regarding the infertility-related anxiety items at baseline, women in the maximal arm were more fearful of definitive childlessness (p=0.033). This difference did not exist in the after-measurement. At after-measurement, women in the minimal arm had significantly higher anxiety scores regarding the effect of stress on the physical relationship with their partner (p=0.013) as well as higher anxiety scores for the presence of a twin pregnancy (p=0.006).

Process evaluation of study interventions

Patients

Of the 696 patients included in the maximal intervention arm in the after-measurement, 260 (37%) unique patients reported to have received one or more of the patient leaflets. The majority of those patients appeared highly satisfied with the leaflets (97%). They reported an increased knowledge of potential causes (71%), treatment procedures (90%) and guidelines (51%), an increased understanding of their doctor's treatment policy (61%), an increased ability to ask questions about the treatment (61%. The scores for an improved communication with the doctor (22%), as well as perceived increased empowerment for decision making during consultations (22%) were lower. In total, 83% would want to receive comparable leaflets in the future and 97% would recommend the leaflets to peers.

Table II; indicator scores at baseline and after-measurement

Indicator	Minimal intervention		Change	Maximal intervention		Change	Δ min-max[a]
	Baseline	After		Baseline	After		
Initial assessment of fertility	% (n)	% (n)	%	% (n)	% (n)	%	2-tailed p-value
The initial fertility assessment should result in both a diagnosis and a prognosis.	45.2 (231)	34.6 (165)	- 10.6	14.1 (197)	26.4 (178)	+12.3	p<0.0001 *
The initial fertility assessment should consist of three parts: semenanalysis, tubal occlusion and cycle-analysis.	32.3 (100)	38.8 (116)	+6.5	37.7 (79)	32.0 (100)	-5.7	NS
Couple's history-taking should cover at least: age of both partners, duration of infertility, type of couple's infertility (primary or secondary).	97.0 (209)	99.4 (178)	+2.4	98.0 (152)	98.0 (148)	0	NS
Woman's physical examination should include assessment of the Body Mass Index.	74.1 (193)	92.6 (176)	+18.5	70.7 (140)	76.6 (141)	+5.9	NS
Life-style advice should be part of the counselling regarding pregnancy-probabilities.	50.8 (65)	75.0 (68)	+25.8	65.3 (49)	68.8 (48)	+3.5	p=0.0436*
Advice concerning *smoking*							
Advice concerning *alcohol*	45.2 (115)	67.8 (121)	+22.6	54.6 (77	68.1 (116	+13.5	NS
Advice concerning *bodyweight*	48.9 (90)	53.9 (78)	+5	52.7 (55)	51.4 (70)	-1.3	NS
Anovulation							
Patients with overweight should, with regard to their fertility-treatment and overall health, be informed of the importance of weight reduction by means of life-style changes.	20.3 (69)	25.6 (78)	+5.3	38.6 (57)	27.9 (43)	-10.7	NS
The ovarian response to hormonal stimulation should be performed by regular transvaginal ultrasound (frequency of 1-3 times / week).	86.1 (72)	96.7 (61)	+10.6	88.1 (59)	93.8 (32)	+5.7	NS
Before starting ovulation induction treatment, the patient should be informed about: specific side-effects of medication, the need for regular intensive follow-up during treatment, the increased risks of multiple pregnancy, hyperstimulation syndrome and spontaneous abortion.	13.8 (109)	19.6 (143)	+5.8	23.7 (93)	21.4 (84)	-2.3	NS

Indicator	Minimal intervention		Change	Maximal intervention		Change	Δ min-max[a]
	Baseline	After		Baseline	After		
Male factor infertility	% (n)	% (n)	%	% (n)	% (n)	%	2-tailed p-value
In case of normospermia (WHO-criteria), semen-analysis should not be repeated, and no additional andrological tests should be performed.	89.8 (48)	72.5 (40)	-17.3	66.7 (30)	51.6 (31)	-15.1	NS
In case of an abnormal semen-analysis (WHO-criteria), the physician should perform: 1. andrological history-taking 2. a physical examination 3. at least one extra semen-analysis.	6.9 (30)	2.5 (40)	-4.4	21.4 (29)	31.0 (29)	+9.6	NS
In case of an idiopathic oligoasthenoterato-zoospermia, no hormones, vitamins or NSAID's should be prescribed to improve semen-quality.	100 (88)	100 (86)		100 (91)	100 (76)		NS
Intra uterine insemination (IUI)							
In case of unexplained infertility, stimulated IUI should not be offered, even though it is associated with higher pregnancy rates than unstimulated IUI, because it carries a risk of multiple pregnancy.	8.3 (89)	15.2 (46)	+6.9	28.0 (49)	17.2 (29)	-10.8	NS
The diagnosis 'cervical factor' is an indication for IUI in the unstimulated cycle.	11.1 (18)	13.3 (15)	+2.2	22.7 (23)	11.8 (17)	-10.9	NS
In case of IUI in the stimulated cycle, ovarian response should be monitored by transvaginal ultrasound.	58.3 (103)	26.5 (68)	-31.8	66.7 (51)	61.0 (59)	-5.7	p= 0.0451*
Each department performing IUI should evaluate their results annually	100 (8)	100 (8)	-	100 (8)	100 (8)	-	NS
Indications for IVF/ICSI							
In case of male infertility with VCM<1x10*6 /ml (before capacitation) there is a direct indication for ICSI-treatment.	13.6 (44)	56.7 (30)	+43.1	24.0 (50)	73.6 (19)	+49.6	NS
In case of unexplained infertility in a woman <36 years, there is an indication for IVF after 3 years of infertility.	61.5 (13)	70.0 (10)	+8.5	83.3 (30)	81.3 (16)	-2.0	NS
The routine use of hCG for luteal support after IVF is not recommended	100 (8)	100 (8)	-	100 (8)	100 (8)	-	NS

Indicator	Minimal intervention		Change	Maximal intervention		Change	Δ min-max[a]
	Baseline	After		Baseline	After		
Embryo Act	% (n)	% (n)	%	% (n)	% (n)	%	2-tailed p-value
During an intake prior to IVF-treatment, the following should be discussed:							
The risks of hyperstimulation, poor response and accompanying consequences, complications such as infection and bleeding, the laboratory procedure and the risk of swapping- and laboratory-mistakes.	13.0 (115)	12.9 (155)	+0.1	14.2 (155)	11.6 (138)	-2.6	NS
The chances of success, pregnancy after SET and DET.	82.4 (125)	90.7 (161)	+8.3	89.1 (175)	88.0 (142)	-1.1	NS
The chances of congenital diseases/malformations, ectopic pregnancy, multiple pregnancy and spontaneous abortion and possible unknown long-term risks.	6.4 (109)	8.1 (149)	+1.7	23.8 (147)	12.1 (132)	-11.7	NS
The emotional aspects of the treatment, how patients can communicate that they are in need of extra emotional support and the existence of the patient association for infertility to share experiences and gain information from the patient's perspective.	41.2 (119)	53.5 (144)	+12.3	52.2 (161)	55.2 (143)	0	NS
How hospital care during treatment is organised and which caretakers are involved in the treatment procedure, how patients should contact the fertility team in case of problems or complaints and how the fertility team can be contacted, especially outside daytime-hours.	66.4 (104)	74.0 (146)	+7.6	70.0 (160)	63.6 (129)	-6.4	p=0.0336 **

[a] = corrected for baseline differences; * = significant difference in favor of the maximal intervention;
** = significant difference in favor of the minimal intervention; NS = non significant

Addendum to table II: Some patient numbers or percentages shown in table II regarding the *baseline* measurement may slightly differ from those reported in table I in chapter 4. Increased insight during the study period did the researchers decide to make minor adjustments in the complex algorithms behind some indicators. This encompassed for example adding stricter exclusion criteria and loosening assessment of history-taking items by including only recommendations with high-grade evidence to the algorithm. Some indicators that concerned more than one treatment cycle per patient, were operationalized anew by formulating indicator adherence as a dichotomized score per patient instead of indicator adherence per cycle separately; this way, the care assessment by indicators approximates daily practice more closer, as professionals may have legitimate reasons to defer from guidelines in specific cases. Adjustments were done to include as many cases as reasonably possible without loosening inclusion criteria or deviating from the guideline recommendation contents. Indicators of both baseline as well as after-measurement in table II were calculated using the exactly the same indicator algorithms, and are thus comparable over time.

Professionals

Professionals from both intervention arms rated the feedback report as easily accessible, comprehensible and reported that the report actually contributed to the implementation of the guidelines in their clinics. In the maximal intervention arm, the feedback meeting was rated of equal value. In those maximal intervention clinics, the 'List of suggested tools for implementation of the guidelines at local level' was moreover highly appreciated, whereas professionals were indifferent to the leaflet on Shared Decision Making and the patient information checklists. One remarkable reason that came up frequently for not using the offered intervention material for fertility guideline implementation, was that it was considered to be "not my job responsibility" to initiate practice changes.

Discussion

To our knowledge, we report the first trial to evaluate the implementation of a set of fertility guidelines, and one of few trials in general that searches to influence professional behavior by a patient oriented intervention. However, both the minimal professional directed strategy and the maximal tailored, multifaceted patient and professional oriented strategy did not improve the overall use of recommendations in our national guidelines on comprehensive fertility care. For only two indicators significant effects were seen in favor of the maximal intervention, compared to the minimal intervention.

Not having found a sustainable effect of either strategy for the entire guideline programme, raises the question whether these specific strategies are ineffective in itself, or whether ineffectiveness is largely caused by flaws in study design or incomplete execution of the intervention strategies. We will further elaborate on these possibilities.

Rationale of the tested interventions

From the reviews of the Cochrane Effective Practice and Organisation of Care Group (EPOC), we learned that the impact of different interventions to promote healthcare interventions is largely variable. Both 'audit and feedback' and 'educational outreach visits' are reported to have a small to modest effect on improving professional practice and the effect of 'printed educational materials' seems to be only beneficial when compared to 'no intervention'.[19] In the field of reproductive healthcare, a recent review of evidence-based strategies for implementing guidelines in obstetrics[42], showed especially positive effects for interventions including audit and feedback, reminders, as well as multifaceted strategies. It also clearly demonstrated that the prospective identification of barriers to change leads to better adaptation of interventions. Recent studies on barriers to implementation in reproductive medicine identified mainly organizational barriers.[23,25,27,29,30,43] As the professionals involved in obstetric, gynaecological and fertility care are largely similar, and evidence for implementation within fertility care is very scarce, an extrapolation of these results may be assumed to be reasonable. However, we have to keep in mind that the major pitfall in implementing change into practice lies therein that previously successful interventions may prove worthless in other settings. This assumption to extrapolate

and rely on previously performed barrier research and interventions, might be one explanation why the tested strategies in our c-RCT did not prove to be successful.

Study Design

We choose to perform a c-RCT to evaluate the effect of the tested interventions, as this is considered the 'gold standard' in implementation research.[44] Randomization was performed at clinic rather than professional or patient level to avoid any risk of contamination of both study arms. However, such a clustered design requests more patients to achieve an appropriate study power, which can be done by either increasing the number of clusters or the number of patients per cluster. However, as the patient sample was randomly selected from electronic databases for financial purposes in which treatment or diagnosis type was not indicated, sufficient inclusion per indicator could not be guaranteed beforehand. Incomplete inclusion unfortunately made it very hard to achieve statistical power for some of the less frequently applicable indicators. Proving changes in effect of the interventions was moreover complicated because effect-sizes in implementation research are in general only small to moderate (5%-15%).[15,21,45,46]

Choosing a complete set of guidelines with a large number of concrete recommendations for practice as a study subject was an ambitious and challenging task. The same professionals are involved in providing diagnosis and treatment of the various types of infertility, so interventions were aimed at these professionals irrespective of background characteristics of their patients (i.e. diagnosis or treatment type). If the interventions would have had an overall positive effect on guideline adherence, a large group of patients could thus have benefitted from an increase in health gain and overall efficacy of care. This in return, would have had a substantial beneficial impact on cost-effectiveness of the strategies. However, taking the lack of observed effects for the comprehensive set of recommendations into account, and the difficulties of including enough patients of some of the indicators, some side remarks should be made for future studies. We recommend to exert a future c-RCT with a more strictly specified subgroup of patients, for example with the shared diagnosis 'male infertility and a more limited set of recommendations'. This would have several advantages. First, the patient sample could be easier identified by either clinical or laboratory registrations. Second, when enrolling the intervention, professionals only have to focus on a distinct patient group and limited clinical care pathways, thus probably increasing their awareness to the provided interventions.

Incomplete execution of the intervention

The patient leaflet was received by only 37% and read by 30% of intended patients from the maximal intervention clinics. The vast majority of leaflet recipients read the material (87%, 92% and 87% respectively for each of the leaflets) and highly appreciated it. We learn from these results that patients on the one hand are amenable to this innovative type of implementation strategy, whereas the professionals on the other hand either appeared to be quite indifferent to the execution of this intervention, or did not consider it their job to do so. Subgroup analysis of patients that actually received the patient information leaflets was unfortunately not possible because of low numbers per indicator. Obviously the

success of a maximal intervention as described in this paper depends on optimal dissemination of intervention materials, for example due to the presence of local opinion-leaders who commit themselves to the project. We moreover recommend to search for alternative ways of implementing such interventions, for example direct automated mailing of material through identification of patients from e.g. EPF's, systematic dissemination on the floor through specialized nurses or, perhaps quite unorthodox, even by bypassing the already busy professionals by entrusting the patients associations alone with this task.

As the feedback report and meeting were generally rated as the most valuable intervention elements by the professionals, it could very well be possible that the surplus value of the maximal implementation strategy was limited. This is in accordance with literature, which recently questions the previously reported advantages of multi-faceted over single interventions.[16] Change needs time, certainly when it concerns effectuating a change in the communication and decision-making process by patient-education and empowerment. It could be a slow process requiring a redefinition of traditional roles and a paradigm shift in doctor-patient interaction. As stated above, the challenge of implementing a comprehensive set of guidelines by addressing all dimensions of fertility care at the same time is a huge one. A longer implementation period (e.g. 12 or even 24 months) might have been more justified to uncover intervention effects, especially as the interventions were introduced just before the regional summer holiday season.

Patient involvement

So far, most evidence from literature on guideline implementation strategies, concerns professional-oriented and to a lesser extent organisational interventions. The patient is seldom directly involved in introducing change and improvement. However, individual preferences of the patient can indeed have major impact on decisions about which health care is delivered.[47,48] Although this c-RCT did not proof any beneficial effect of either of the tested strategies, we think that the role of patients in influencing and improving health care provision should be a continued focus of attention in future research on fertility care. Negative effects of the interventions, such as an increase of anxiety or depression scores, were not observed. Patients from the maximal intervention arm were even found to be less anxious for twin pregnancy or the effect on the physical relationship with their partner, which could be caused by more open communication with their care providers. Regarding the infertility-related anxiety items, women in the maximal arm were more fearful of definitive childlessness ($p=0.033$) at baseline than at the after-measurement. This could be caused by positive ascertainment from their doctors our from increased understanding of their problem due to the patient leaflets, however, it might also be explained by longer duration of infertility and higher frequency of primary infertility at baseline.

Summarizing, the results of our study, although not evidently effective, can contribute to an increased understanding of the potential role of patients in clinical guideline implementation, as the process evaluation data on the patient oriented intervention showed promising results; patients did feel empowered to act as a partner in the diagnostic and treatment process and experienced an improvement in communication.

Whether professionals are also prepared to accept patients as equal partners in clinical decision-making, remains another challenging focus for further research.[49] Similarly, the role of leadership, team climate, and local 'culture for change' could be interesting subjects for future research[50,51], particularly because some professionals indeed felt and reported that 'initiating practice changes' was not their responsibility.

In conclusion, in this c-RCT we did not find an overall sustainable effect of the studied professional as well as patient oriented interventions, which aimed to improve the use of a set of national guidelines for comprehensive fertility care. However, an evaluation of the patient oriented intervention showed generally positive experiences with patient empowerment and doctor-patient communication. Professionals appeared quite indifferent to the disseminated implementation materials, raising the question whether they should be the only actors responsible for guideline implementation. This provides a promising basis for further studies on the effects of patient oriented implementation of professional guidelines.

References

1. Vayena E, Rowe PJ,Griffin P. Current practices and controversies in assisted reproduction: report of a WHO meeting. Geneva: World Health Organisation, 2002.
2. Nachtigall RD. International disparities in access to infertility services. Fertil Steril 2006;85: 871-5.
3. Evers JL. Female subfertility. Lancet 2002;360:151-9.
4. Boivin J, Bunting L, Collins JA, Nygren KG. International estimates of infertility prevalence and treatment-seeking: potential need and demand for infertility medical care. Hum Reprod 2007;22:1506-12.
5. Oakley L, Doyle P, Maconochie N. Lifetime prevalence of infertility and infertility treatment in the UK: results from a population-based survey of reproduction. Hum Reprod 2008;23:447-50.
6. Bhattacharya S, Porter M, Amalraj E, Templeton A, Hamilton M, Lee AJ, Kurinczuk JJ. The epidemiology of infertility in the North East of Scotland. Hum Reprod 2009.
7. Greil AL. Infertility and psychological distress: a critical review of the literature. Soc Sci Med 1997;45:1679-1704.
8. Kallen B, Finnstrom O, Nygren KG, Otterblad OP, Wennerholm UB. In vitro fertilisation in Sweden: obstetric characteristics, maternal morbidity and mortality. BJOG 2005;112:1529-35.
9. Ombelet W, Martens G, De SP, Gerris J, Bosmans E, Ruyssinck G, Defoort P, Molenberghs G, Gyselaers W. Perinatal outcome of 12,021 singleton and 3108 twin births after non-IVF-assisted reproduction: a cohort study. Hum Reprod 2006;21:1025-32.
10. Cousineau TM, Domar AD. Psychological impact of infertility. Best Pract Res Clin Obstet Gynaecol 2007; in press.
11. Schieve LA, Cohen B, Nannini A, Ferre C, Reynolds MA, Zhang Z, Jeng G, Macaluso M, Wright VC. A population-based study of maternal and perinatal outcomes associated with assisted reproductive technology in Massachusetts. Matern Child Health J 2007;11:517-25.
12. Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing physician performance. A systematic review of the effect of continuing medical education strategies. JAMA 1995;274:700-5.
13. Wensing M, van der Weijden T, Grol R. Implementing guidelines and innovations in general practice: which interventions are effective? Br J Gen Pract 1998;48:991-7.
14. Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. Lancet 2003;362:1225-30.
15. Grimshaw J, Eccles M, Tetroe J. Implementing clinical guidelines: current evidence and future implications. J Contin Educ Health Prof 2004;24(Suppl 1):S31-S37.
16. Grimshaw J, Eccles M, Thomas R, Maclennan G, Ramsay C, Fraser C, Vale L. Toward evidence-based quality improvement. Evidence (and its limitations) of the effectiveness of guideline dissemination and implementation strategies 1966-1998. J Gen Intern Med 2006;21(Suppl 2):S14-S20.
17. Jamtvedt G, Young JM, Kristoffersen DT, O'Brien MA, Oxman AD. Audit and feedback: effects on professional practice and health care outcomes. Cochrane Database Syst Rev CD000259, 2006.
18. O'Brien MA, Rogers S, Jamtvedt G, Oxman AD, Odgaard-Jensen J, Kristoffersen DT, Forsetlund L, Bainbridge D, Freemantle N, Davis DA, et al. Educational outreach visits: effects on professional practice and health care outcomes. Cochrane Database Syst Rev CD000409, 2007.
19. Farmer AP, Legare F, Turcot L, Grimshaw J, Harvey E, McGowan JL, Wolf F. Printed educational materials: effects on professional practice and health care outcomes. Cochrane Database Syst Rev CD004398, 2008.
20. Griffin SJ, Kinmonth AL, Veltman MW, Gillard S, Grant J, Stewart M. Effect on health-related outcomes of interventions to alter the interaction between patients and practitioners: a systematic review of trials. Ann Fam Med 2004;2:595-608.
21. Grol R, Wensing M, Eccles M. Improving Patient Care; the implementation of change in clinical practice.London: Elsevier Butterworth Heinemann, 2005.
22. Wetzels R, Harmsen M, van Weel C, Grol R, Wensing M. Interventions for improving older patients' involvement in primary care episodes. Cochrane Database Syst Rev CD004273, 2007.
23. Haagen EC, Nelen WL, Hermens RP, Braat DD, Grol RP, Kremer JA. Barriers to physician adherence to a subfertility guideline. Hum Reprod 2005;20:3301-6.
24. Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PA, Rubin HR. Why don't physicians follow clinical practice guidelines? A framework for improvement. JAMA 1999;282:1458-65.

25. Cardenas V, Davis RL, Hasselquist MB, Zavitkovsky A and Schuchat A. Barriers to implementing the group B streptococcal prevention guidelines. Birth 2002;29:285-290.
26. Foy R, Maclennan G, Grimshaw J, Penney G, Campbell M, Grol R. Attributes of clinical recommendations that influence change in practice following audit and feedback. J Clin Epidemiol 2002;55:717-22.
27. Farquhar CM, Kofa E, Power ML, Zinberg S, Schulkin J. Clinical practice guidelines as educational tools for obstetrician-gynecologists. J Reprod Med 2002;47:897-902.
28. Graham ID, Logan J, Davies B, Nimrod C. Changing the use of electronic fetal monitoring and labor support: a case study of barriers and facilitators. Birth 2004;31:293-301.
29. Foy R, Walker A, Ramsay C, Penney G, Grimshaw J, Francis J. Theory-based identification of barriers to quality improvement: induced abortion care. Int J Qual Health Care 2005;17:147-55.
30. Say L, Foy R. Improving induced abortion care in Scotland: enablers and constraints. J Fam Plann Reprod Health Care 2005;31:20-3.
31. Chan JK, Kapp DS, Shin JY, Osann K, Leiserowitz GS, Cress RD, O'Malley C. Factors associated with the suboptimal treatment of women less than 55 years of age with early-stage ovarian cancer. Gynecol Oncol 2008;108:95-9.
32. Shaw B, Cheater F, Baker R, Gillies C, Hearnshaw H, Flottorp S, Robertson N. Tailored interventions to overcome identified barriers to change: effects on professional practice and health care outcomes. Cochrane Database Syst Rev CD005470, 2005.
33. Campbell MK, Elbourne DR, Altman DG. CONSORT statement: extension to cluster randomised trials. BMJ 2004;328:702-8.
34. Mourad SM, Hermens RP, Nelen WL, Braat DD, Grol RP, Kremer JA. Guideline-based development of quality indicators for subfertility care. Hum Reprod 2007;22:2665-72.
35. Mourad SM, Nelen WL, Hermens RP, Bancsi LF, Braat DD, Zielhuis GA, Grol RP, Kremer JA. Variation in subfertility care measured by guideline-based performance indicators. Hum Reprod 2008;23:2493-500.
36. Mourad SM, Hermens RP, Cox-Witbraad T, Grol RP, Nelen WL, Kremer JA. Information provision in fertility care: a call for improvement. Hum Reprod 2009;24(6):1420-6.
37. Landis JR and Koch GG The measurement of observer agreement for categorical data. Biometrics 1977; 33, 159-174.
38. Spielberger C., Gorsuch R, & Lushene, R. (1970). Manual for the State-Trait Anxiety Inventory. Palo Alto, CA: Consulting Psychologist Press
39. Van Der Ploeg H, Defares P and Spielberger C Handleiding bij de Zelf-Beoordelings Vragenlijst ZBV: Een nederlandstalige bewerking van de Spielberger State-Trait Anxiety Inventory, STAI-DY (1980). Swets & Zeitlinger B.V., Lisse, The Netherlands
40. Beck AT, Guth D, Steer RA and Ball R Screening for major depression disorders in medical inpatients with the Beck Depression Inventory for Primary Care. Behav Res Ther 1997; 35, 785-791.
41. Hulscher ME, Laurant MG, Grol RP. Process evaluation on quality improvement interventions. Qual Saf Health Care 2003;12:40-6.
42. Chaillet N, Dube E, Dugas M, Audibert F, Tourigny C, Fraser WD, Dumont A. Evidence-based strategies for implementing guidelines in obstetrics: a systematic review. Obstet Gynecol 2006;108:1234-45.
43. Graham ID, Logan J, Davies B, Nimrod C. Changing the use of electronic fetal monitoring and labor support: a case study of barriers and facilitators. Birth 2004;31:293-301.
44. Ovretveit J, Gustafson D. Evaluation of quality improvement programmes. Qual Saf Health Care 2002;11:270-5.
45. Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD, Thomson MA. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. The Cochrane Effective Practice and Organization of Care Review Group. BMJ 1998;317:465-8.
46. Grol R, Grimshaw J. Evidence-based implementation of evidence-based medicine. Jt Comm J Qual Improv 1999;25:503-13.
47. Wensing M, Elwyn G. Methods for incorporating patients' views in health care. BMJ 2003;326:877-9.

48. Elwyn G, Hutchings H, Edwards A, Rapport F, Wensing M, Cheung WY, Grol R. The OPTION scale: measuring the extent that clinicians involve patients in decision-making tasks. Health Expect 2005;8:34-42.
49. Jung HP, Wensing M, Olesen F, Grol R. Comparison of patients' and general practitioners' evaluations of general practice care. Qual Saf Health Care 2002;11:315-9.
50. Davies HT, Nutley SM, Mannion R. Organisational culture and quality of health care. Qual Health Care 2000;9:111-9.
51. Bower P, Campbell S, Bojke C, Sibbald B. Team structure, team climate and the quality of care in primary care: an observational study. Qual Saf Health Care 2003;12:273-9.

Chapter 8

General discussion

Summarizing

This thesis focused on improving the implementation of a set of clinical practice guidelines on comprehensive fertility care.

First, a set of guideline-based quality indicators was developed by means of a rigorous systematic Delphi-method. Such quality indicators are important to assess current care and, by repeated assessment, also to monitor changes in care. The indicator-set was extensively tested in a practice test based on medical records and questionnaire data, to explore its value as an instrument within a quality improvement initiative. Additional professional and patient questionnaires were used to get more insight into several specific aspects of current fertility care: A determinant study was performed to gain knowledge on potential patient and clinic characteristics associated with the current practice of information provision. A second determinant study investigated the potential patient and clinic characteristics that are related to patient satisfaction and experiences with current fertility care. Subsequently, a cluster-randomized controlled trial (c-RCT) was performed in which an innovative multi-faceted, tailored, professional and patient oriented strategy was tested against a single-faceted professional oriented strategy to implement the Dutch fertility guideline programme. The practice test was used as a baseline measurement in this c-RCT. Finally, a process evaluation was performed to gain insight into the implementation process and degree of effectiveness of both strategies.

In this final chapter, we present answers to the research questions posed in the general introduction. The main findings from the studies in the previous chapters and some methodological considerations are discussed in the light of available literature. The chapter will conclude with some implications for future research as well as future practice.

Answers to the research questions

The following answers to the research questions posed in chapter 1, can be formulated from the studies described in this thesis.

1. **What can be regarded as a valid set of quality indicators for comprehensive fertility care? (chapter 2, 3, 4)**

   Chapter two describes the rationale for the types of indicators, i.e. process and structure, that were sought after in the presented study in chapter 3. Following a systematic, rigorous six-step Delphi method, we were able to develop a guideline-based set of 39 process and structure indicators for comprehensive fertility care, both evidence and consensus-based. In a practice test, seven quality criteria per indicator were evaluated in chapter 4 to explore the validity and feasibility of these indicators as instruments to assess clinical care. Those criteria were: measurability, applicability, improvement potential, discriminatory capacity, complexity and case-mix stability. The proposed indicator set scored variably on the different quality criteria; overall, the set proved suitable for assessing and improving comprehensive fertility care. However, tailored to a similarly sized and selected patient sample as was used in the described practice test, we could clearly distinguish three subsets within the original set: a first subset of indicators with high applicability and

improvement potential proved suitable for monitoring and quality improvement purposes; a second subset of indicators with low improvement potential to be used for quality monitoring purposes only; lastly, a third subset of indicators which requires a different, very specific fertility patient-sample to make them useful for quality improvement purposes.

2. **What is the quality of current fertility care in the Netherlands, and how is this experienced by patients? (chapter 4, 5, 6)**
   Based on the extensive audit of comprehensive fertility care by means of the described set of quality indicators in chapter 4, a large variation between sites in the quality of current fertility care in the Netherlands was found. For some indicators, the variation between the clinics, ranged up to a 100%. The median adherence to guideline recommendations was less than 50% for 14 of the 39 indicators assessed. Although 100% scores may not be realistic due to practical problems in indicator measurement and the need to divert from guideline recommendations in specific cases, all clinics should strive to equal the highest scoring clinic, which is a goal within reach. From this current care assessment, it can be concluded that there is ample room for improvement within Dutch clinical fertility care. This is particularly true for the practice of information provision, of which we performed an extensive case study in chapter 5, including a determinants analysis. It was shown that an alarming number of couples is deprived of sufficient or complete information regarding diagnostics, treatment course and concurrent risks and complications. However, a huge discrepancy between these findings and the way current information provision was experienced by patients was found. The vast majority of patients reported that the information they received was sufficient. In addition, chapter 6 showed that patients' experiences with fertility care were predominantly positive and satisfaction with provided care was high. This was partly associated with specific patient characteristics like high education level, achieving pregnancy or good mental health. Moreover, both determinant studies showed that small and sometimes simple changes in the organization of current care, e.g. the use of checklists for information provision, the presence of trained fertility nurses or a specialized fertility consultation hour, could effectuate higher quality and higher rated fertility care.

3. **Does a combined professional and patient oriented intervention strategy prove more effective for the implementation of a fertility guideline programme than an exclusively professional oriented strategy? (chapter 7)**
   The cluster-randomized controlled trial described in chapter 7, revealed that neither the innovative multi-faceted professional and patient oriented strategy nor the single professional oriented strategy improved the actual use of the comprehensive Dutch fertility guidelines as a whole. On the one hand, within the subset of indicators with high improvement potential, no overall improvement was measured. On the other hand, within the subset of indicators with already high adherence (i.e. indicators with low improvement potential) it could be seen that care remained of comparably high standards.

4. **What are successful elements of the intervention strategies used? (chapter 7)**
   A process-evaluation of the c-RCT was performed among professionals and patients, in order to shed some light on the use and appreciation of the implementation tools. Professionals from both intervention arms rated the feedback report as easily accessible, comprehensible and reported that the report actually contributed most to the implementation of the guidelines in their clinics. In the maximal intervention arm, the feedback meeting and 'List of suggested tools for implementation of the guidelines at local level' were moreover appreciated, whereas professionals were indifferent to the remaining intervention elements. Patients were generally positive regarding the distributed leaflets. They reported an increase in knowledge of potential causes of the fertility problem, fertility treatment options and fertility guidelines. Moreover, they perceived an increased understanding of their doctor's treatment policy as well as increased empowerment to influence treatment decisions during consultations. However, as distribution of the patient-leaflets was poor, other methods of introduction should be sought-after for this kind of intervention to be successful.

Discussion of main findings

The studies performed within the span of this thesis taught us that an objective care assessment by means of systematically developed quality indicators can identify the domains of fertility care in need of improvement. Both professionals and healthcare authorities are thus enabled to act upon this gathered knowledge on current care performance. Improvement strategies can be developed which are tailored to the different care settings and local shortcomings in performance can be addressed.

Assessing and improving fertility care

Defining high quality of fertility care

The first step in each quality improvement initiative should be the definition of appropriate outcome measures. Commonly used traditional measures to evaluate the quality of care, are mortality and morbidity. For fertility care, the most frequently used outcome measures are 'pregnancy rate' or 'life birth rate', and to a much lesser extent morbidity, reflected in complication rate. However, a high pregnancy rate does not automatically mean that the quality of delivered care is high. For instance, we do not consider fertility care to be of high quality when professionals achieve high pregnancy rates at the cost of an increased number of potential life-threatening complications like OHSS or multiple pregnancies. Moreover, as discussed in chapter 2 and 3, the traditional outcome measures do not provide us with information on where exactly improvements in care are necessary. Measures that focus on the process and structure of care could indeed provide such information. However, an universal definition of what exactly should be "high quality fertility care" does not exist yet. Pennings and Ombelet proposed for fertility care that all treatment regimens should refer to the standards of "patient-friendly ART" [1]. Van Empel and colleagues redefined this concept to "high quality ART", while entrusting the patient with a key position regarding all six aspects of high quality care of the Institute of Medicine (IOM): safety, effectiveness, patient-centeredness, timeliness, efficiency and equity.[2,3] Following the quality initiative proposed by the IOM, high quality care should thus be measured by each of

these six dimensions. Although professionals seem increasingly supportive of the central position that patients should have in fertility care, it will require a paradigm shift in professionals' sense of quality to replace, or rather complete, the traditional outcome measures with these new dimensions. We therefore promote the use of a broad set of indicators covering several aspects of care, instead of the rather short-sighted use of only traditional outcomes measures. As existing national guidelines represent a condensed version of widely available scientific evidence and professional consensus, they offer a good basis for the definition of such indicators.

Developing quality indicators for fertility care

The specific aim of the study described in chapter 3, was the development of a valid indicator set suitable for the assessment of comprehensive clinical fertility care. By involving end-users like gynaecologists and fertility clinicians as panel-members in the development-process, a solid basis for future implementation activities was created. A systematic procedure was performed conform the current 'state of the art' in indicator development.[4,5] Firstly, a literature search for existing indicator sets was performed and the Dutch fertility guideline programme was taken as a starting point for fertility indicator development. The literature search showed the rather surprising finding that an international equivalent of this indicator-development project was not yet performed for clinical fertility care. An initiative that showed the closest resemblance, was the proposed set of seven 'key priorities for implementation' accompanying the NICE fertility guideline which was issued by the National Health Service (NHS).[6] However, these key priorities were not systematically developed, let alone tested before guideline dissemination. The priority topics were explicitly chosen because of convenience of data collection by the requested use of existing NHS resources. This lack of precedents within clinical fertility care makes the indicator-studies presented in this thesis unique to the field and assigns them an explorative and maybe even exemplary status.

Our entire indicator development process was performed bearing the specific goal of assessing comprehensive clinical fertility care in mind. This does however implicate that concessions have been made regarding the total selected number of indicators per guideline, to avoid ending up with a far too large indicator set. Such concessions are however inherent to any selection procedure. Our six-step method resulted in a very well 'weighted' choice of those guideline recommendations that were considered most important regarding overall efficacy, health gain and improvement potential for the sole purpose of assessing comprehensive fertility care. If however a need is felt for better or more extensive monitoring of one of the separate guideline topics, a new indicator procedure can, and should be, executed per guideline. As the indicator set was primarily developed for internal audit purposes in a research setting, external uptake and use of the set or widespread publication of results should be guided carefully, as misinterpretation and potential misuse by uneducated media could easily occur.[7,8]

The status of guideline-based indicators in quality of care assessments

There are some important issues to keep in mind when using and interpreting quality indicators for care assessments, even if they are robustly developed. Foremost, it is necessary to realize that quality indicators do not provide a black and white picture of assessed care in absolute terms of what's wrong and what's good. The scope of the described indicator assessment was primarily the identification of potential shortcomings in current care. This does however not mean that shortcomings (i.e. non-adherence to guideline recommendations) are automatically of clinical significance. Therefore, imperfect quality indicator scores need to be seen as a signal that delivered care should be evaluated more closely, rather than a statement that delivered care is insufficient per se.[4]

For guideline-based quality indicators in particular, there is another important point of attention. In daily practice, there might be good reasons for professionals to divert from a guideline, for example when specific treatment circumstances or individual patients' profile require so. Here, we come to a frequently heard objection (i.e. from the professional's point of view) to guideline-based quality indicator scores. Guidelines are developed for the vast majority of patients, and guideline adherence can thus be hampered by a professional's need to tend to an individual patient's situation. Such 'motivated exceptions' are very hard to register or take into account in performance assessments, as they are not necessarily written down in medical records, nor systematically entered in the concise entries within electronic databases. In our current care assessments, we tried to take into account the potentially motivated exceptions as much as possible, for example, by calculating per patient a weighed indicator score that is computed from, for instance, the sum of various cycles. An elaborated example can be given on the calculation of the indicator recommending unstimulated IUI in case of idiopathic infertility. It is imaginable that a need for a different treatment policy can be felt after performing already four cycles according to the guideline recommendation that did not result into a pregnancy. In this case, a certain degree of leniency is allowed to the professional regarding a well-considered use of ovarian stimulation in the fifth and sixth cycles. The aberrance from the guideline recommendation in only two out of 6 cycles will thus not directly influence the weighed indicator score, as two-thirds of the cycles were still according to the guideline-recommendation.

Guideline and quality indicator dynamics

Guidelines, as well as accompanying quality indicators should be treated as 'living' concepts and should therefore be periodically updated. Some of the studied guidelines in this thesis were unfortunately quite aged, with a maximum age of nine years for the guidelines 'male infertility', 'OHSS' and 'indications for IVF' at the time of the after-measurement. For example, in case of idiopathic infertility, the Dutch IUI guideline (issued in 2000), prescribes stimulated IUI as treatment of choice. In contrast, the NICE guideline issued by the NHS in 2004 recommends unstimulated IUI to prevent adverse outcomes like multiple pregnancies, although there is evidence stating that stimulated IUI has better outcomes in terms of pregnancy rates. Dutch fertility clinicians are individually adapting to this recommendation that became available in a newly spread

NHS best-practice guideline, by switching their policy from stimulated to unstimulated IUI. We anticipated upon this by enabling panel members to add new recommendations in the first step of the Delphi procedure. We should however bear in mind that the most important prerequisite for delivering evidence-based best care, is that evidence should become widely available to professionals, by timely or continuous updates.

Implementing change into current practice

In general, results of controlled trials and systematic reviews on guideline implementation show that efforts are often not very successful. If any improvement is achieved at all, these are small-to-moderate (around 5%-10%); however, such changes may still be clinically relevant.[9-12] The impact on patient outcomes has often not been studied at all or could not be evidently demonstrated.[13,14] The evaluated interventions in the c-RCT described in chapter 7 proved unfortunately not very beneficial in improving current fertility care. This of course raises the question about underlying causes. We will discuss here as potential causes the currently available evidence of implementation techniques, available knowledge of barriers to implementation and the actual extent of implementation of the executed strategies.

The Cochrane Effective Practice and Organisation of Care Group (EPOC) aims to systematically review the impact of different interventions to promote healthcare interventions. Both 'Audit and feedback' and 'educational outreach visits' are reported to have a small to modest effect on improving professional practice.[15,16] The effect of 'printed educational materials' (PEM) seems to be beneficial when compared to no intervention, although this effectiveness is uncertain when PEM's are compared to other interventions.[17] Cameron et al. investigated how feedback on clinical performance of different intensities (i.e. feedback, feedback plus action planning letter, feedback plus facilitated action planning) was received by maternity professionals[18]. Although feedback of increased intensity proved feasible and acceptable to clinicians, the authors were unable to demonstrate that clinicians actually increased their intention to comply with audit criteria. It thus remains unclear whether more intensive strategies should be preferred above simple ones. One of the factors that can intuitively be depicted as a contributor to the degree of success of the introduced implementation strategies, lies within the local level of 'quality improvement attitude', meaning the organizational culture or the team's willingness to change. This potential contributor was however not defined as subject for research in this thesis, whereas we can predict that the presence of a favourable quality improvement culture is an important factor contributing to successful guideline implementation. Moreover, we did not evaluate the presence and impact of local opinion-leaders who introduce and promote implementation policy, although the presence of such a person is previously reported to be important.[19] In our c-RCT, professionals in the maximal intervention arm had to open up to the idea that patients have knowledge of, and could actually utilise professional guidelines to their advantage. If patients in general adapted to this, it would definitely change traditional doctor-patient roles and interaction. So, opinion leaders could play a substantial role in guiding colleagues through this change in doctor and patient culture. As for now, the

interesting question remains whether both patient as well as doctors are ready yet to accept this shift towards a more active patient role in discussing and planning treatment policy. The evidence-base for effects of having a 'change culture' or 'change promoter' remains unfortunately incomplete and in this field also, further research is necessary.[20-23]

For reproductive medicine, quite some efforts have recently been made to expand the knowledge on guideline adherence and guideline implementation. The first sporadic publications date from the mid-nineties, with a steady surge in the number of publications on several guideline topics from the year 2000 onwards. [18,24-34] New research projects are increasingly carried out and as a result the number of publications is growing as well as the evidence base.

However, for fertility care in particular, evidence on guideline implementation is still lacking. One of the core problems of implementation research in general, is that what proves to be effective in one setting, may not work at all in another setting. Therefore, while formulating interventions for guideline implementation in fertility care, researchers should be very cautious in extrapolating results from publications regarding neighboring fields of reproductive medicine, for instance obstetrics. This does not mean however, that this knowledge can't be used as a starting point for new trials in fertility care, as in many settings the same gynaecologists are active in both obstetric, gynaecologic as well as fertility care.

A review of evidence-based strategies for implementing guidelines in obstetrics[35], showed especially positive effects for interventions including audit and feedback, reminders, as well as multifaceted strategies. It also clearly demonstrated that the prospective identification of barriers to change leads to better adaptation of interventions. As previous evidence concerning this barrier identification remained unequivocal[36], this obstetric review offers an important basis for future studies among reproductive health care professionals. A newly published study on barriers among gynaecologists dealing with induced abortion care[37], identified mainly organizational barriers, which is not very different from other studies in reproductive medicine[25,27,38-40].

In summary, literature remains largely inconclusive about the extent of effectiveness of the different interventions applied in our c-RCT; moreover, the effectiveness of interventions that were of value in one setting, cannot be guaranteed when translated to other settings. The lack of any beneficial effect of the interventions in the described c-RCT could thus be caused by an imperfect theoretical background, the lack of a barrier analysis performed uniquely for this project, or simply inappropriateness of the interventions for translation to our particular study setting.

Process-evaluation

A process evaluation was performed to gain more insight into the potential causes of ineffectiveness of the interventions by increasing our knowledge of the events during the intervention period. For the evaluation of any interventions aimed at healthcare improvement, performing a rigorous process evaluation is at least as important as a (c-)RCT of interventions, as it clearly shows where exactly an effect was established or failed to occur. Although the coverage rate of the patient-oriented intervention was imperfect (37% of included patients received one or more leaflets), the evaluation of

the leaflets by patients was predominantly positive. Due to the leaflets, patients reported an increased knowledge of potential causes, treatment and guidelines, an increased understanding of their doctor's treatment policy as well as perceived increased empowerment for decision making during consultations. Professionals, however, appeared quite indifferent to the disseminated implementation materials. It could thus very well be possible that the surplus value of the maximal implementation strategy was limited when compared to the minimal strategy. This also raises the question whether professionals should be the only actors responsible for guideline implementation; future efforts to induce practice changes might prove more efficient by concentrating on patients or patient associations.

Methodological considerations

Study design

Performing a (c-)RCT to evaluate interventions for healthcare improvement, is considered the 'gold standard' in implementation research.[41] We chose a clustered design with a randomization at clinic level because randomization at patient or professional level would be practically impossible without risking contamination of both study arms. However, such a design requests many more patients to achieve an appropriate study power. This can be effectuated by increasing the number of clusters or the number of patients per cluster. As we randomly selected our patient sample from an electronic financial database in which treatment or diagnosis was not indicated, we could not guarantee sufficient inclusion per indicator beforehand. Including more patients or more clusters in order to try to achieve higher statistical power, would have considerably decreased overall feasibility of the study, as data collection was already an extensive task.

In this thesis, we aimed to assess and improve the quality of comprehensive clinical fertility care according to a national guideline programme consisting of ten guidelines, in 16 regional clinics. This accounted for an ambitious multi-centre study focusing on a complete guideline programme without a precursor within reproductive medicine. There is one major advantage of trying to implement a complete guideline programme at once. Some of the known barriers for fertility guideline implementation, for example 'lack of self-efficacy' [27], are not unique to one guideline. As the same fertility professionals are involved for each of the guidelines, the implementation of the entire programme can thus be improved by removing this barriers with a single intervention. For example, increasing the degree of shared decision-making to overcome the 'lack of self-efficacy', can be used for both patients within the 'initial assessment of fertility' who turn down an expectative treatment policy, but is also applicable to patients that want to pursue a second IVF-cycle even if less than three follicles have grown after maximal ovarian hyper stimulation. Introducing an intervention for all guidelines at once, could thus prove very cost-effective because different subgroups of patients and thus different aspects of fertility care could be improved with a single intervention at professional level.

Study population

Conversely, improvement of so many issues at the same time was probably too difficult.

Neither the minimal nor the maximal implementation strategy had an evident beneficial effect on the scores of indicators deducted from the ten national guidelines on comprehensive fertility care. Shrinking scale size, for example to only one guideline or only one subgroup of patients (for instance, only IUI or IVF/ICSI patients) receiving care from multiple guidelines, could have some benefits. Not only would it relieve practical problems in indicator measurement (e.g. making it less time consuming to identify and assess cases) but professionals would also be provided with a clearer focus in adapting and using the offered interventions, as one clinical problem or patient group could then have their undivided 'quality improvement attention'. An additional advantage is that sufficient power for correct statistical analysis of results will be more easily achieved if the patient sample is more homogenous in terms of diagnosis or treatment type. In brief, achieving optimal results in comparable study settings, could be achieved by confinement and aggravation of implementation efforts by focusing more on details.

The role of patients in guideline implementation

The rationale to involve patients in fertility guideline implementation derives from the hypothesis that fertility patients, who are a relatively young, critical and actively participating patient group, have high interests in influencing and understanding their treatment policy and might also favor a high degree of shared decision making. Griffin and colleagues performed a review on health-related outcomes of interventions intended to alter the interaction between patients and professionals [42]. They concluded that patient targeted approaches were more effective than those targeted at professionals. Moreover, providing patients with specific prompt sheets proved more effective than providing only generalized information, even without changing the length of consultations. Literature from social science reports that opportunities for involvement can increase patients' self-confidence, self esteem, and level of social contact[43]. For maternity care, the effect of coaching patients to participate in clinical decision making and the use of question prompts for patients, lead to increased perception of control and increased involvement in care[44]: a need for exactly these two concepts is a frequently heard preference from focus group interviews among fertility patients from the participating clinics. Therefore, it was hypothesized that if any patient group could thus be able to influence health care delivery by direct interaction with their caregiver, it could very well be fertility patients.

Taking the previous considerations, cited literature and process-evaluation data into account, we can conclude that the patient-directed implementation strategy developed for our c-RCT could be successful among this patient sample in theory, but remains hampered by the limited amount of evidence within implementation research. Moreover, if the indicators targeted were for example restricted to a single guideline, or if the patient sample would have been more homogenous, this patient-directed intervention could have been more elaborated and tailored to those groups. Now, we can only conclude that more research is needed to broaden the evidence

base for patient-oriented guideline implementation interventions in the field of fertility care.

Practical execution

From a practical point of view, assessing ten guidelines in one project, proved a laborious exercise. Even though participating clinics belonged to one collaborating and internally referring region, different registration standards and methods for medical data existed at every location, for example different types of electronic or paper charts and different levels of completeness of treatment registration. A lot of effort was put into gathering medical records at the several locations and training the extra personnel (three 5th year medical students and one health care researcher) for data extraction. Moreover, practical problems in indicator operationalization could not be fully anticipated upon in the study design phase. Poor availability of medical data and the lack of adequate data resources are common problems in performance assessment efforts[45,46] .

Electronic patient files

However, difficulties to properly measure clinical care, or more specifically quality indicators, should never be a reason to discard rigorously selected indicators. On the contrary, efforts should be undertaken to facilitate measurement initiatives. The introduction of a national Electronic Patient File (EPF) in the near future, literally offers a blank page for those interested in clinical performance assessments. Professionals and professional societies alike should therefore realize that it's time to stop lingering and being only a suspicious observer of EPF-progression; the time has come to actively join the EPF-movement and seize this great opportunity for future continuous monitoring of actual care. In particular, attention should be paid to relatively small investments or adjustments to existing databases that can be made for this purpose. The routinely monitoring of clinical performance indicators, such as our developed set for fertility care, within an EPF will greatly increase our knowledge of current care as well as of the practice of indicator assessment.

Level of analysis

Data were gathered and analysed at clinic level and data on the performance of individual professionals was therefore not available[37]. This was done because performance data were often only traceable at team level: Sometimes individual signatures lack in medical charts or, in particular for teaching clinics, treatment policy should be attributed to a supervisor's or joint team decision. In chapter 6 we therefore included only hospital characteristics like care infrastructure (e.g. presence of a separate fertility consultation hour) and team composition (e.g. presence of trained fertility nurses) as potential determinants at the care provider's level. Previous barrier-research, however, showed that there are also barriers to guideline adherence at the individual professional's level[27]. Assessing indicator scores at the individual professional's level would thus provide information on performance that is the closest to the care process. However, to disentangle each individual's contribution to a complex care performances implicates major methodological challenges: not only is a considerable number of patients needed per professional for adequate statistical

analysis of such 'individualized indicator scores', but attributing each policy decision to a traceable individual would be practically utopian.

Ceiling effect

A lack of performance improvement after an intervention might be attributed to a failure of the intervention itself or a failure to measure change when levels of performance are already high at baseline[15]. It is possible that, due to motivated exceptions to guideline recommendations or practical measurement difficulties, any change in the fourth quartile of adherence (75%-100%) is harder to achieve and assess. Moreover, there could be a 'ceiling effect' due to age of the guidelines. It is likely that the first period after new evidence or guidelines become available is critical for adaptation of the contents and thus for changing current care. This implicates that adherence to guideline recommendations that is not achieved within the first months or years following guideline dissemination, is a priori unlikely to happen anymore. An explanation could be that at least part of the non-adherence can be attributed to a professional's disagreement with guideline contents and thus unwillingness to change current practice. Summarising, chances of successful implementation are likely to decrease with an increasing interval of first publication.

Implications for future research and practice

From the above, several implications can be distilled for future research as well as future practice.

Implications for future research on fertility guideline implementation

There are some methodological recommendations that emerge from our experience with the studies performed within this thesis. These involve the following topics:

STUDY DESIGN

Considering a c-RCT design with a formal control group and multiple intervention arms, instead of two intervention arms and a historic control group, could give more insight into effect of interventions of diverse intensity. An additional advantage would be that such a study design lends itself better for a cost-effectiveness comparison of the different interventions. However, more clusters or patients are needed for such a design, which might decrease feasibility of data-collection.

STUDY OBJECT

As discussed before, some of the studied guidelines were relatively aged at time of the implementation effort. Picking a soon-to-be or newly ratified guideline as subject to a follow-up study, could offer a more promising departure point for renewed evaluation of (some of) the proposed implementation strategies. This would offer the possibility to perform a baseline assessment of current care just before dissemination of the guideline, immediately followed by a (cluster randomized controlled) trial of implementation policies of different intensity. Aiming to implement such a recently issued instead of an aged guideline will (at least partly) prevent discussion on validity of guideline contents or status of underlying evidence by the involved professionals.

Study instruments

Including patients in indicator development will contribute to patient-centeredness of the quality of care assessment. However, as patients frequently indicate problems with the assessment of medical-technical recommendations, or a lack of self-efficacy in a panel with professionals, a separate and pre-trained patient panel should be considered. Furthermore, it might be necessary to appoint an extra scoring weight to those patient-selected recommendations during the overall final indicator selection round, to guarantee a high degree of uptake of those patient-selected recommendations for indicator-selection.

Collecting data on patients' health care needs is considered to lead to better detection of short-comings in care, but we demonstrated that there exists a vast discrepancy between patient experiences and satisfaction when assessed by means of questionnaires on the one hand, and objective care assessment by means of quality indicators on the other hand. The two types of assessment should therefore always be combined, to reflect the state of current care as completely as possible.

Study outcomes

Within the scope and design of our project it was unfortunately not possible to make a proper outcome assessment regarding pregnancy rates, multiple pregnancy rates or life birth rates. Because of the heterogeneity of the patient sample, pregnancy rates could not be validly compared between patients; for some patients only had one consultation or went through an initial assessment of fertility, whereas other patients might have had three consecutive embryo transfers in the study period. This means that endpoints in time were often not comparable between subgroups of patients. A study design in which a single patient subgroup is subject to study, for example, IUI or ovulation induction patients lends itself more properly for such comparisons. Linking structure and process measures from the indicator set to patient outcomes like pregnancy rate, would provide an interesting insight into the complex relationship between the two[47].

Theoretical background

The above mentioned suggestions for further research aim more or less at a more perfect study design to test the proposed interventions. It might, however, very well be possible that our focus should be on entirely different interventions. More research should for example be conducted to unravel the role of local culture for change, team climate and opinion leaders in improving guideline implementation.

Implications for future policy and practice in fertility care

Based on the topics described in this thesis and discussed in this last chapter, some implications can be formulated for policy and current practice. We will present these implications while addressing the relevant stakeholders involved:

IMPLICATIONS FOR POLICY OF PROFESSIONAL ORGANIZATIONS

- Supplying guidelines with sets of quality indicators and creating the (electronic) infrastructure to measure them, can help professionals to identify those domains of care in need of improvement and moreover enables continuous quality monitoring. Letting professionals participate in the development process, creates transparency and a supporting base for future implementation activities.
- An important prerequisite for the successful use of clinical practice guidelines or sets of quality indicators is that both are kept 'alive', meaning continuously updated with newly available evidence.
- Convenient assessment of quality indicators should be facilitated by the uptake of those indicators in the future nationwide EPF.
- The effort pays back; guidelines and quality indicators can be tools to achieve high quality care. Inability or difficulty to measure an indicator should never be the sole reason to abandon this indicator or relent from assessment efforts.
- The use of large, (inter-)national registrations enable solid assessment of less frequent conditions and complications by means of quality indicators.
- To achieve high quality care, the major efforts generally put into guideline development, should be equalled by efforts put into guideline implementation activities and guideline implementation research.
- Providing patients with leaflets on professional guideline contents can increase their understanding professional decision-making, improve doctor-patient communication and empower patients to influence treatment decisions. All this can improve a doctor's perception of self-efficacy and thus lead to better guideline adherence.
- Professional organisations should stress the importance of shared decision making, and encourage their residents-in-training to implement this concept.

IMPLICATIONS FOR PROFESSIONALS

- Participating in current care assessments by means of quality indicators is very useful as it provides insight into those aspects of current care where local improvements are needed.
- Professionals should think of those assessments as an opportunity, not as merely another bureaucratic interference. Assessment results can educate and guide towards optimal health care delivery.
- Guidelines are as a rule formulated for 'the majority' of patients. Making motivated exceptions to these guidelines for individual patients should always be possible, but should not result from a lack of self-efficacy concerning doctor-patient interaction.
- In fertility care, the current practice of information provision is insufficient. Whether caused by limited information provision by professionals or defective memorization by patients, in both cases professionals should take responsibility to

improve current care. The presence of specialized nursing personnel or the use of information checklists by professionals could improve this care.
- As information provision for highly educated, Dutch couples seems already insufficient, special attention should be paid to risk-groups that are lowly educated or non-Dutch.

IMPLICATIONS FOR PATIENTS AND PATIENT ASSOCIATIONS
- Taking notice of professional guideline contents, either by specialized leaflets or visiting the professional society's website, can increase patients' understanding of their doctor's decision-making, improve communication with the doctor and empower the patient to influence treatment decisions.
- Although occurrence of complications is fortunately infrequent, make sure patients are optimally informed about the potential risks and complications of future treatment.
- Having an active and well-organised patient association can provide the structure for information-exchange between patient and professional stakeholders, for example, by clarifying information about professional guidelines.
- Encouraging patients to take up their role in shared decision-making, can trigger medical professionals to apply it more often.
- It will probably prove beneficial if patient oriented interventions are not introduced by professionals but instead by patient organisations or representatives.

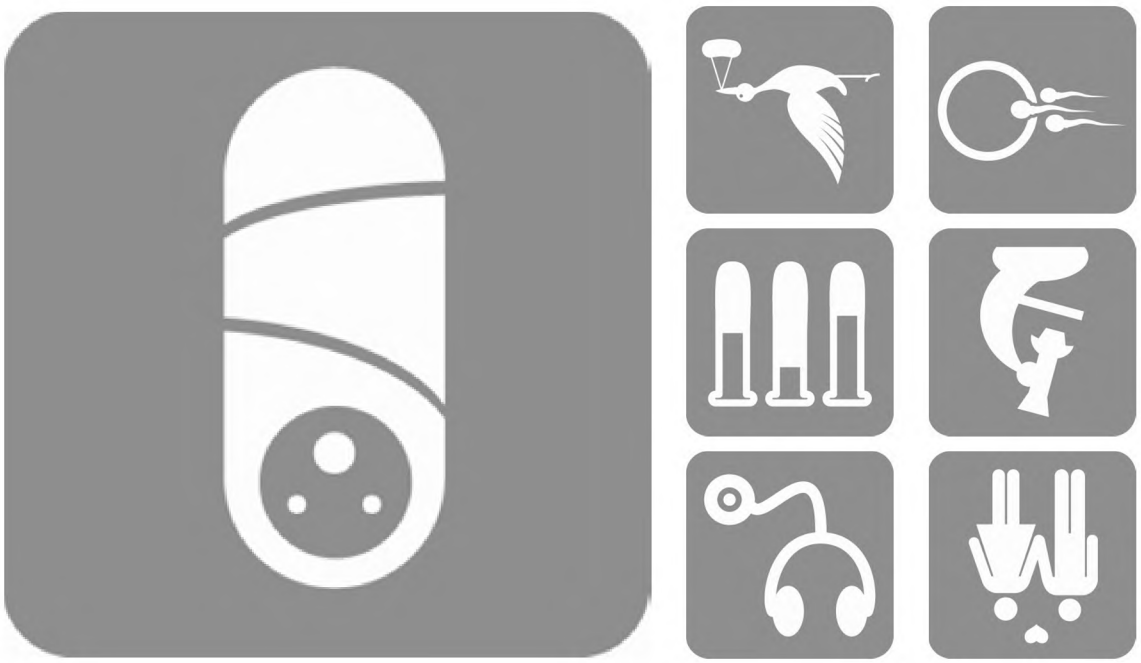
Final conclusion

In conclusion, we developed a valid indicator set that can be used for monitoring and improving fertility care. We could not demonstrate an universal benefit of the tested strategies for the implementation of the Dutch fertility guidelines in a c-RCT. However, the process evaluation among patients of the maximal implementation group, showed promising results on the use of patient leaflets explaining professional guideline contents, especially regarding understanding of the professional's treatment policy and empowerment for participation in clinical decision-making. These are ingredients to entrust patients with a central role within guideline implementation in the future. However, challenges for further research on the improvement of the quality of current fertility care by guideline implementation are manifold, and should be simultaneously taken up by the different stakeholders involved.

References

1. Pennings G, Ombelet W. Coming soon to your clinic: patient-friendly ART. Hum Reprod 2007;22:2075-9.
2. Institute of Medicine. Crossing the quality chasm: a new health system for the twenty-first century. 2001.
3. van Empel I, Nelen WL, Hermens RP, Kremer JA. Coming soon to your clinic: high-quality ART. Hum Reprod 2008;23:1242-5.
4. Campbell SM, Braspenning J, Hutchinson A, Marshall MN. Research methods used in developing and applying quality indicators in primary care. BMJ 2003;326:816-9.
5. Wollersheim H, Hermens R, Hulscher M, Braspenning J, Ouwens M, Schouten J, Marres H, Dijkstra R, Grol R. Clinical indicators: development and applications. Neth J Med 2007;65:15-22.
6. NICE. Fertility: assessment and treatment for people with fertility problems (http://www.nice.org.uk/nicemedia/pdf/CG011fullguideline.pdf). 2004.
7. Lilford R, Mohammed MA, Spiegelhalter D, Thomson R. Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. Lancet 2004;363:1147-54.
8. Hannan EL, Chassin MR. Publicly reporting quality information. JAMA 2005;293:2999-3000.
9. Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD, Thomson MA. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. The Cochrane Effective Practice and Organization of Care Review Group. BMJ 1998;317:465-8.
10. Grol R, Grimshaw J. Evidence-based implementation of evidence-based medicine. Jt Comm J Qual Improv 1999;25:503-13.
11. Grimshaw J, Eccles M, Tetroe J. Implementing clinical guidelines: current evidence and future implications. J Contin Educ Health Prof 2004;24(Suppl 1):S31-S37.
12. Grol R, Wensing M, Eccles M. Improving Patient Care; the implementation of change in clinical practice. London: Elsevier Butterworth Heinemann, 2005.
13. Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. Lancet 1993;342:1317-22.
14. Grol R. Improving the quality of medical care: building bridges among professional pride, payer profit, and patient satisfaction. JAMA 2001;286:2578-85.
15. Jamtvedt G, Young JM, Kristoffersen DT, O'Brien MA, Oxman AD. Audit and feedback: effects on professional practice and health care outcomes. Cochrane Database Syst Rev CD000259, 2006.
16. O'Brien MA, Rogers S, Jamtvedt G, Oxman AD, Odgaard-Jensen J, Kristoffersen DT, Forsetlund L, Bainbridge D, Freemantle N, Davis DA, et al. Educational outreach visits: effects on professional practice and health care outcomes. Cochrane Database Syst Rev CD000409, 2007.
17. Farmer AP, Legare F, Turcot L, Grimshaw J, Harvey E, McGowan JL, Wolf F. Printed educational materials: effects on professional practice and health care outcomes. Cochrane Database Syst Rev CD004398, 2008.
18. Cameron M, Penney G, Maclennan G, McLeer S, Walker A. Impact on maternity professionals of novel approaches to clinical audit feedback. Eval Health Prof 2007;30:75-95.
19. Doumit G, Gattellari M, Grimshaw J, O'Brien MA. Local opinion leaders: effects on professional practice and health care outcomes. Cochrane Database Syst Rev CD000125, 2007.
20. Soumerai SB, McLaughlin TJ, Gurwitz JH, Guadagnoli E, Hauptman PJ, Borbas C, Morris N, McLaughlin B, Gao X, Willison DJ, et al. Effect of local medical opinion leaders on quality of care for acute myocardial infarction: a randomized controlled trial. JAMA 1998;279:1358-63.
21. Borbas C, Morris N, McLaughlin B, Asinger R, Gobel F. The role of clinical opinion leaders in guideline implementation and quality improvement. Chest 2000;118:24S-32S.
22. Berner ES, Baker CS, Funkhouser E, Heudebert GR, Allison JJ, Fargason CA, Jr., Li Q, Person SD, Kiefe CI. Do local opinion leaders augment hospital quality improvement efforts? A randomized trial to promote adherence to unstable angina guidelines. Med Care 2003;41:420-31.
23. Bosch M, Dijkstra R, Wensing M, van der Weijden T, Grol R. Organizational culture, team climate and diabetes care in small office-based practices. BMC Health Serv Res 2008;8:180.
24. Mousa HA, Mahmood TA. Do practice guidelines guide practice? A prospective audit of induction of labor three years experience. Acta Obstet Gynecol Scand 2000;79:1086-92.

25. Graham ID, Logan J, Davies B, Nimrod C. Changing the use of electronic fetal monitoring and labor support: a case study of barriers and facilitators. Birth 2004;31:293-301.
26. Chapin J, Root W. Improving obstetrician-gynecologist implementation of smoking cessation guidelines for pregnant women: an interim report of the American College of Obstetricians and Gynecologists. Nicotine Tob Res 2004;6(Suppl 2):S253-S257.
27. Haagen EC, Nelen WL, Hermens RP, Braat DD, Grol RP, Kremer JA. Barriers to physician adherence to a subfertility guideline. Hum Reprod 2005;20:3301-6.
28. Haagen EC, Hermens RP, Nelen WL, Braat DD, Grol RP, Kremer JA. Subfertility guidelines in Europe: the quantity and quality of intrauterine insemination guidelines. Hum Reprod 2006;21:2103-9.
29. Jordan TR, Dake JR, Price JH. Best practices for smoking cessation in pregnancy: do obstetrician/gynecologists use them in practice? J Womens Health (Larchmt) 2006;15:400-41.
30. Curtis M, Abelman S, Schulkin J, Williams JL, Fassett EM. Do we practice what we preach? A review of actual clinical practice with regards to preconception care guidelines. Matern Child Health J 2006;10:S53-S58.
31. Gulmezoglu AM, Langer A, Piaggio G, Lumbiganon P, Villar J, Grimshaw J. Cluster randomised trial of an active, multifaceted educational intervention based on the WHO Reproductive Health Library to improve obstetric practices. BJOG 2007;114:16-23.
32. Stern JE, Cedars MI, Jain T, Klein NA, Beaird CM, Grainger DA, Gibbons WE. Assisted reproductive technology practice patterns and the impact of embryo transfer guidelines in the United States. Fertil Steril 2007;88:275-82.
33. Franssen MT, Korevaar JC, van d, V, Boer K, Leschot NJ, Goddijn M. Management of recurrent miscarriage: evaluating the impact of a guideline. Hum Reprod 2007;22:1298-303.
34. Brand PL, van Lingen RA. [Guideline for imminent preterm birth: insufficient adherence and implementation]. Ned Tijdschr Geneeskd 2008;152:359-61.
35. Chaillet N, Dube E, Dugas M, Audibert F, Tourigny C, Fraser WD, Dumont A. Evidence-based strategies for implementing guidelines in obstetrics: a systematic review. Obstet Gynecol 2006;108:1234-45.
36. Shaw B, Cheater F, Baker R, Gillies C, Hearnshaw H, Flottorp S, Robertson N. Tailored interventions to overcome identified barriers to change: effects on professional practice and health care outcomes. Cochrane Database Syst Rev CD005470, 2005.
37. Foy R, Walker A, Ramsay C, Penney G, Grimshaw J, Francis J. Theory-based identification of barriers to quality improvement: induced abortion care. Int J Qual Health Care 2005;17:147-55.
38. Cardenas V, Davis RL, Hasselquist MB, Zavitkovsky A, Schuchat A. Barriers to implementing the group B streptococcal prevention guidelines. Birth 2002;29:285-90.
39. Farquhar CM, Kofa E, Power ML, Zinberg S, Schulkin J. Clinical practice guidelines as educational tools for obstetrician-gynecologists. J Reprod Med 2002;47:897-902.
40. Say L, Foy R. Improving induced abortion care in Scotland: enablers and constraints. J Fam Plann Reprod Health Care 2005;31:20-3.
41. Ovretveit J, Gustafson D. Evaluation of quality improvement programmes. Qual Saf Health Care 2002;11:270-5.
42. Griffin SJ, Kinmonth AL, Veltman MW, Gillard S, Grant J, Stewart M. Effect on health-related outcomes of interventions to alter the interaction between patients and practitioners: a systematic review of trials. Ann Fam Med 2004;2:595-608.
43. Carr S. Has service user participation made a difference to social care services? Social Care Institute for Excellence, 2002.
44. Rowe RE, Garcia J, Macfarlane AJ, Davidson LL. Improving communication between health professionals and women in maternity care: a structured review. Health Expect 2002;5:63-83.
45. Gagliardi AR, Fung MF, Langer B, Stern H, Brown AD. Development of ovarian cancer surgery quality indicators using a modified Delphi approach. Gynecol Oncol 2005;97:446-56.
46. Pronovost PJ, Miller M, Wachter RM. The GAAP in quality measurement and reporting. JAMA 2007;298:1800-2.
47. Rubin HR, Pronovost P, Diette GB. The advantages and disadvantages of process-based measures of health care quality. Int J Qual Health Care 2001;13:469-74.



Chapter 9

Summary

Samenvatting

Summary

Chapter 1: General Introduction

Chapter 1 describes the rationale for the studies performed within this thesis. Infertility is a worldwide problem with high societal impact. To help clinicians provide evidence and consensus based healthcare, diverse (inter)national organizations have sought to develop guidelines for the provision of fertility care. In the Netherlands, the Dutch Society for Obstetrics and Gynaecology (NVOG) has developed a broad guideline programme to guide fertility professionals in providing their daily care. However, as guidelines do not implement themselves, we describe the background and current state of evidence for guideline implementation in general. While evidence for fertility guideline implementation is currently lacking, we deduct from the available evidence and barrier analyses within general and reproductive medicine two experimental strategies to improve implementation of the Dutch fertility guideline programme. As fertility patients are in general young and critical about their care, we hypothesize they could be just the allies we need in implementing fertility guidelines.

Chapter 2: Debate on outcome measures for fertility care

Monitoring fertility care by the use of quality indicators will facilitate the transparency of provided health care as well as comparisons over time and between settings. However, what exactly should be the best indicators to assess fertility care is still subject to debate. We contribute to this discussion by suggesting the systematic development of evidence-based quality indicators, especially process and structure indicators, rather than outcome measures, for the monitoring and improvement of fertility care.

Chapter 3: Indicator development for fertility care

As none of the existing fertility guideline programmes worldwide is accompanied by a satisfactory set of quality indicators, we aimed in chapter 3 to develop a valid, guideline-based indicator set. A systematic RAND-modified Delphi method was used to develop a set of key recommendations based on 10 national Dutch fertility guidelines, international literature and existing international indicators. Experts' opinions were used to appraise recommendations regarding specific criteria such as efficacy, level of health gain, applicability and potential for care improvement. As a result, a representative set of 39 key recommendations was selected from 303 initial recommendations. The recommendations covered two structural and 37 procedural aspects, the latter encompassing 'indications for treatment', 'diagnostic procedures', 'treatment procedures' and 'patient information'. These recommendations were, as a

final step, translated into measureable indicators for monitoring and improving clinical fertility care.

Chapter 4: Variation in fertility care

In chapter 4, a practice test was performed to test the previously developed indicator set. This study aimed to assess several quality criteria of the indicator-set and to use the set to assess current fertility care. For this purpose, a historic cohort study was performed in 16 Dutch fertility clinics. In total, 2698 couples were invited to participate. Indicator data were gathered by medical record extraction, and patient and professional questionnaires. Quality criteria for each indicator (measurability, reliability, applicability, improvement potential, discriminatory capacity, complexity and case-mix stability) were assessed. Current practice was measured as each clinic's adherence to the separate indicators. As a result, 1499 (56%) couples participated. All indicators were measurable, but the results for the other quality criteria varied. In total, 14 of the 39 indicators scored <50% adherence. There existed a wide variation in performance between the clinics, ranging up to 100%. The highest median adherence (86%) was found within the guideline 'indications for IVF-treatment'. The lowest median adherence is found within the guideline 'initial assessment of fertility' (43%), followed closely by the guideline 'anovulation' (44%). Based on the results of our practice test, particularly applicability and improvement potential, we can distinguish three subsets within our original set of 39 quality indicators. Indicators within the first subset score highly for applicability and improvement potential and can therefore be used for quality monitoring purposes and as a baseline measurement for quality improvement programmes in the involved clinics. A second subset contains indicators with low improvement potential, which makes them useless for improvement programmes, but with high applicability, which makes them still suitable for monitoring purposes, namely to ensure that adherence continues to be high in the future. Within the third subset we find indicators with low applicability. They concern relatively rare conditions or complications (e.g. premature ovarian failure and severe ovarian hyperstimulation syndrome), so inclusion numbers within our random patient sample were too small for adequate indicator assessment. In conclusion, this study shows that the quality of the developed indicator-set was adequate for the monitoring of clinical fertility care. Such objective assessment of care can help professionals to identify and subsequently target the domains of care in need of improvement A first assessment in the Netherlands reveals large variation between clinics and ample room for improvement of care.

Chapter 5: Information provision in fertility care

In chapter 5, we focused on the current practice of information provision in fertility care, as adequate information provision is a crucial dimension of high-quality care. Clinical practice guidelines containing consensus-based recommendations may standardize practice between settings. The described study was designed for three purposes: (i) to assess actual adherence to recommendations on information provision, (ii) to measure patient satisfaction with current practice and (iii) to analyse how variation in adherence relates to the characteristics of patients and clinics. To do so, all recommendations concerning patient information were extracted from the national fertility guideline programme and edited into a patient questionnaire. Additional questions concerning patient satisfaction and potential determinants of information provision at patient level were included (e.g. female age, type of infertility). A professional's questionnaire was sent to all gynaecologists to gather potential determinants at clinic level (e.g. fertility team size, presence of specialized fertility nurses). Multilevel regression analysis was performed to identify the determinants of information provision. A total of 1499 couples (56%) participated. The percentage of couples who reported to have received complete information varied between recommendations from 10 to 96% (mean 57%). Overall, 94% of couples were satisfied with fertility services. The use of checklists for information provision, the presence of obstetrics/gynaecology residents and specialized nursing personnel, and higher patient anxiety scores were significantly associated ($P < 0.05$) with higher levels of information received. We conclude from these results that information provision for infertile couples is currently poor and in need of improvement. This could easily be procured by, for example, the use of information checklists.

Chapter 6: Patients' experiences with fertility care

Chapter 6 aimed to asses patients' experiences and satisfaction with fertility care as well as the determinants associated with these two concepts. The questionnaire that was sent to the previously described patient sample included specific questions for this purpose.

Waiting times, information provision and emotional support were experienced the least positive aspects of care. Determinants of all care aspects were found significant ($p<0.05$) at four different domains. Three at patient level, i.e. demographic characteristics, type of received treatment and both general and mental health status, and one at clinic level, i.e. organization of care. In general, patients' satisfaction with care was rated high (94%).

The results from this study provide us with an increased understanding of the determinants of patients' experiences and satisfaction with fertility care. This enables

professionals to tailor their care to specific subgroups of patients and adjust their organization of fertility care where needed. Moreover, the study underlines the need to investigate whether case-mix correction is necessary whenever interpreting patient-surveys on care experiences, as both patient's and clinic's characteristics can influence the way healthcare delivery is experienced.

Chapter 7: A cluster-randomized controlled trial to improve fertility guideline implementation

Chapter 7 describes a cluster-randomized controlled trial designed to evaluate the effect of two different strategies to improve the implementation of the Dutch fertility guideline programme. A single, professional oriented strategy of audit and feedback was tested against a multifaceted, tailored professional and patient oriented strategy consisting of audit and feedback, dissemination of professional implementation material and specific patient leaflets that explain the professional guideline contents in lay language. The used outcome measures were the extent of adherence to the indicators developed in chapter 3. The data from chapter 3 was taken as a before measurement. Neither strategy provided an overall beneficial effect on the implementation of the fertility guideline programme. Small effects were seen for isolated indicators. However, patient evaluation of the patient leaflets was very positive, with reported positive effects on doctor-patient communication and an experienced increased empowerment for participation in the clinical decision making process. A possible explanation of the lack of effects is the poor dissemination of intervention material among patients and poor use of the intervention elements by participating professionals. The surplus value of the maximal intervention can be questioned as the most highly appreciated intervention elements were the feedback report and feedback meeting.

Chapter 8: General discussion

In the final chapter, we present the main findings of the studies from the previous chapters. The research questions from the introductory chapter are answered and discussed in the light of available literature. The developed set of indicators from chapter 3 seems to be a valid instrument for monitoring and improving fertility care. However, we might need to use a differently chosen patient sample or convert to subgroups of indicators for the different purposes of monitoring and improving care. Methodological considerations of the performed studies are taken into account while interpreting these results. The lack of an overall effect of the maximal intervention does not rule out a role for patients within professional guideline implementation, as

patient evaluation of the strategy showed promising results concerning communication and empowerment.

However, trying to improve an entire guideline programme for comprehensive fertility care proved an ambitious task. Suggestions for future research on patient oriented guideline implementation are therefore described. As an overall effect of the strategies was not found, future research might profit from a different study design in which only a subgroup of patients or a single, recently issued guideline is the focus of attention. In addition, more research should be conducted to augment our theoretic knowledge of for example the role of local culture for change, team climate and opinion leaders in improving guideline implementation.

Moreover, practice and policy implications for healthcare providers as well as patients are therefore in this final chapter. We state for example that a convenient assessment of quality indicators should be facilitated by the uptake of those indicators in the future nationwide EPF. Similarly, the use of large, (inter-)national registrations could enable solid assessment of less frequent conditions and complications by means of quality indicators. And finally, the major efforts generally put into guideline development, should be equaled by efforts put into guideline implementation activities and guideline implementation research.

Samenvatting

Hoofdstuk 1: Algemene inleiding

Onvruchtbaarheid (Infertiliteit) wordt omschreven als het uitblijven van zwangerschap na 12 maanden regelmatige, onbeschermde geslachtsgemeenschap. Uit studies is bekend dat wereldwijd bij 4 tot 30% van de paren in de vruchtbare leeftijd sprake is van infertiliteit. Dit betekent dat ongeveer 80 miljoen paren ter wereld te kampen hebben met dit probleem. Het is dan ook niet vreemd dat infertiliteit als één van de grote gezondheidsproblemen van de 21e eeuw wordt beschouwd. Omdat een aanzienlijk deel van de bevolking er vroeg of laat mee te kampen krijgt, heeft Infertiliteit als gezondheidsprobleem dan ook grote impact op onze huidige maatschappij. Niet alleen wat betreft persoonlijk verdriet en eventuele psychische problemen die het oplevert, maar ook omdat de vraag naar nader medisch onderzoek of behandeling groot is. Om behandelaren te helpen zorg te leveren die kwalitatief gezien zo hoogwaardig mogelijk is (dat wil zeggen zorg die is gebaseerd op de meest recente kennis en wetenschappelijke bewijzen of bij gebrek daaraan op door behandelaren onderling overeengekomen aanbevelingen), zijn er diverse professionele organisaties die richtlijnen voor onderzoek naar of behandeling van fertiliteitsproblemen hebben opgesteld. De Nederlandse Vereniging voor Obstetrie en Gynaecologie (NVOG) heeft voor Nederland een negental richtlijnen opgesteld; daarbij komt de nationale Embryo-wet, waarin de zorg rondom IVF en ICSI behandelingen omschreven staat. Door het gebruik van deze richtlijnen kan de variatie in de zorg tussen verschillende behandelaren en verschillende locaties zo klein mogelijk gehouden worden. Helaas is het hebben van richtlijnen alleen, niet genoeg om goede zorg te garanderen; de richtlijnen moeten ook in de dagelijkse praktijk worden ingevoerd, het zogenaamde implementeren. We beschrijven in hoofdstuk 1 de kennis die momenteel beschikbaar is over het implementeren van richtlijnen in het algemeen en in de voortplantingsgeneeskunde in het bijzonder. Ondanks het feit dat de kennis over laatstgenoemde slechts spaarzaam beschikbaar is, ontwikkelen we op basis van de schaarse kennis die beschikbaar is, twee strategieën waarmee we willen proberen de implementatie van de Nederlandse fertiliteitsrichtlijnen te bevorderen. Aangezien fertiliteitspatiënten in het algemeen vrij jong en kritisch zijn, is het heel goed mogelijk dat zij de cruciale bondgenoot kunnen zijn bij het implementeren van fertiliteitsrichtlijnen .

Hoofdstuk 2: Debat over uitkomstmaten binnen fertiliteitszorg

Het hebben van gedegen kwaliteitsindicatoren als meetinstrumenten om fertilliteitszorg te monitoren, zou als voordeel hebben dat de geleverde zorg inzichtelijk gemaakt wordt. Bovendien stelt het ons in staat zorg die op verschillende tijdstippen en plaatsen geleverd wordt, met elkaar te vergelijken. Desalniettemin is er nog discussie wat precies de meest geschikte indicatoren zouden zijn om dergelijke zorg te beoordelen. In hoofdstuk 2 dragen wij bij aan deze internationaal gevoerde discussie. Wij stellen voor dat er op systematische wijze en zo veel mogelijk op basis van wetenschappelijke bewijzen kwaliteits-indicatoren ontwikkeld dienen te worden, die niet alleen op uitkomst gericht zijn (zoals bijvoorbeeld zwangerschap of levendgeborenen per gestarte IVF cyclus) maar met name ook op de proces en structuur van de geleverde zorg (zoals bijvoorbeeld correcte informatievoorziening respectievelijk certificering van het laboratorium).

Hoofdstuk 3: Het ontwikkelen van kwaliteitsindicatoren voor fertiliteitszorg

Aangezien geen van de bestaande internationale richtlijnprogramma's met betrekking op fertiliteitsproblemen, voorzien is van een set van geschikte kwaliteitsindicatoren, hebben wij in hoofdstuk 3 geprobeerd een valide indicatorenset te ontwikkelen die op de richtlijnen gebaseerd is. Een zogenaamde RAND-gemodificeerde Delphi methode werd toegepast om een set van kern-aanbevelingen samen te stellen gebaseerd op de Nederlandse richtlijnen, internationaal beschikbare literatuur en bestaande internationale indicatoren. Aan circa 60 nationale experts werd gevraagd om alle aanbevelingen uit de richtlijnen te beoordelen op specifieke criteria zoals belang voor doelmatigheid van zorg, mate van gezondheidswinst, toepasbaarheid en verbeterbaarheid. Door toepassen van deze procedure werd uit een totaal van 303 aanbevelingen, een definitieve set van 39 kernaanbevelingen geselecteerd. Hiervan hadden er 2 betrekking op de structuur en 37 op procesmaten van zorg. Geselecteerde procesmaten hadden betrekking op de onderwerpen indicatiestelling, diagnostiek, behandeling en informatievoorziening. Als laatste stap werden deze kern-aanbevelingen via algoritmen vertaald naar meetbare indicatoren voor het monitoren en verbeteren van fertiliteitszorg.

Hoofdstuk 4: Variatie in fertiliteitszorg

Met de ontwikkelde indicatorenset uit het voorgaande hoofdstuk, werd in hoofdstuk 4 een uitgebreide praktijktest uitgevoerd. Hierbij werden ook enkele kwaliteitscriteria van de ontwikkelde indicatoren gemeten. Een historisch cohort van 2698 paren fertiliteitspatiënten uit 16 Nederlandse klinieken werd uitgenodigd mee te doen aan deze praktijktest. Data om de indicatoren te meten werd verzameld uit vragenlijsten

naar zowel patiënten en behandelaren als ook uit medische dossiers. Per indicator werden de volgende kwaliteitscriteria gemeten: meetbaarheid, reproduceerbaarheid, toepasbaarheid, verbeterpotentieel, onder-scheidend vermogen, complexiteit en de noodzaak tot correctie voor case-mix. Huidige zorg werd omschreven als de adherentie binnen een kliniek aan de voorgestelde indicatoren. In totaal waren 1499 paren (56%) bereid mee te doen. Hoewel alle indicatoren meetbaar bleken, varieerden de scores op de andere criteria behoorljik. Van de 39 indicatoren werd er voor 14 stuks minder dan 50% adherentie gemeten. Bovendien was er een brede variatie in adherentie tussen de klinieken, namelijk tot wel 100%. De richtlijn 'indicaties voor IVF' scoorde het beste met gemiddeld 86% adherentie, in tegenstelling tot bijvoorbeeld de richtlijn 'orienterend fertiliteitsonderzoek' (43%) en 'anovulatie (44%).

Op basis van de gemeten kwaliteitscriteria kunnen we de oorspronkelijke indicatorenset indelen in drie subgroepen. Een eerste groep indicatoren scoort hoog op toepasbaarheid en verbeterpotentieel, waarmee zij uitermate geschikt is voor zowel monitoring op verschillende momenten in tijd als voor het initiëren van verbeterprogramma's. Dit in tegenstelling tot een tweede groep, met weliswaar hoge toepasbaarheid maar een laag verbeterpotentieel, waardoor die groep ongeschikt is voor verbeterprogramma's maar wel voor continue monitoring van zorg. Tot slot een derde subgroep met slechts een lage toepasbaarheid in de hier onderzochte patiëntensteekproef, aangezien het minder frequent voorkomende klinische beelden of complicaties betreft, zoals prematuur ovarieel falen of een ovarieel overstimulatie syndroom. Om deze indicatorengroep te kunnen meten zou een veel specifiekere patiënteninclusie nodig zijn. Concluderend toont deze studie aan dat de initiële indicatorenset adequaat is voor het monitoren van klinische fertiliteitszorg. Bovendien blijkt er in Nederland sprake van een grote variatie in geleverde fertiliteitszorg tussen verschillende klinieken. Een objectieve meting van de huidige zorg, zoals in deze praktijktest beschreven, kan behandelaren helpen om die onderdelen van zorg te identificeren die verbetering behoeven.

Hoofdstuk 5: Informatievoorziening aan fertiliteitspatiënten

In hoofdstuk 5 hebben we ons toegespitst op de specifieke gang van zaken rond informatievoorziening aan paren die een fertiliteitstraject doorlopen, aangezien goede informatie een cruciale vereiste is om zorg van hoge kwaliteit te noemen. De in dit hoofdstuk beschreven studie concentreerde zich op 3 doelen:

1) het meten van de huidige zorg, dat wil zeggen de mate van adherentie aan richtlijn-aanbevelingen die betrekking hebben op patiënteninformatie
2) het meten van patiënttevredenheid met deze huidige zorg

3) het onderzoeken of variatie in richtlijn-adherentie wellicht te maken heeft met bepaalde karakteristieken van de patiënten of kliniek

Voor dit doel werden alle aanbevelingen met betrekking op informatievoorziening uit de richtlijnen geselecteerd en bewerkt tot een patiëntenvragenlijst. Bovendien werden vragen gericht op tevredenheid en potentiële voorspellers op patiëntniveau (zoals bjivoorbeeld leeftijd van de vrouw) toegevoegd. Daarnaast werd een vragenlijst verzonden aan behandelaren, om potentiële voorspellers op kliniekniveau te verzamelen (zoals bijvoorbeeld de omvang van het behandelteam). Door middel van een multilevel regressie analyse werd geprobeerd voorspellers van kwalitatief goede informatievoorziening te ontdekken. In totaal deden 1499 paren mee. Per informatie-aanbeveling verschilde het aantal paren dat volledige informatie ontving van 10-96% (gemiddeld 56%). Toch was 94% van de paren tevreden met de geleverde zorg. Het bleek dat het gebruik van checklists voor het geven van informatie, de aanwezigheid van gynaecologen in opleiding en het hebben van specifiek geschoold verpleegkundig personeel, evenals angstigere patiënten, geassocieerd waren met het ontvangen van complete informatie. Op basis van deze resultaten concluderen wij dat de huidige Nederlandse zorg met betrekking op informatievoorziening aan fertiliteitspatiënten tekortschiet en voor verbetering vatbaar is. Dit zou gemakkelijk kunnen worden bewerkstelligd door bijvoorbeeld systematische checklists te gebruiken.

Hoofdstuk 6: Ervaringen van patiënten met fertiliteitszorg

De studie in hoofdstuk 6 was erop gericht om enerzijds de ervaringen en tevredenheid van patiënten met fertiliteitszorg te meten, en anderzijds de voorspellers van deze twee begrippen te beschrijven. Dit gebeurde op basis van de reeds eerder beschreven vragenlijst. Met betrekking tot wachttijden, informatie-voorziening en emotionele ondersteuning werden de minst positieve patiëntenervaringen gemeten. Voorspellers van deze ervaringen konden worden gevonden op 4 verschillende domeinen. Drie domeinen hadden betrekking op het niveau van de patiënt, te weten: demografische kenmerken, het type ontvangen behandeling en de gezondheidsstatus (zowel algemeen als mentaal). Het vierde domein werd gevonden op kliniek niveau, en omvatte de organisatie van zorg. Opvallend was, dat de tevredenheid van patiënten met betrekking tot de ontvangen fertiliteitszorg desondanks hoog was (94%).

De resultaten van deze studie vergroten onze kennis van de voorspellers van patiëntenervaringen en -tevredenheid. Deze opgedane kennis stelt behandelaren in staat om hun zorg toe te snijden op specifieke subgroepen van patiënten, en bovendien hun organisatie van zorg waar nodig aan te passen. Tot slot onderschrijft deze studie de noodzaak om correctie van case-mix toe te passen wanneer men

patiëntenenquêtes evalueert, aangezien zowel patiënt als kliniek karakteristieken de wijze waarop geleverde zorg ervaren wordt, kan beïnvloeden.

Hoofdstuk 7: Een cluster-gerandomiseerde gecontroleerde trial om de implementatie van fertiliteitsrichtlijnen te bevorderen

Hoofdstuk 7 beschrijft een clustergerandomiseerde gecontroleerde trial die gericht was op het evalueren van twee verschillende strategieën om de implementatie van het Nederlands richtlijnprogramma van de NVOG met betrekking tot fertiliteits-problemen te verbeteren. Een minimale, enkelvoudige en behandelaargerichte strategie van audit en feedback (het meten van zorg en het terugkoppelen van de resultaten van die meting) werd vergeleken met een maximale, meervoudige strategie die naast op behandelaren, ook op patiënten gericht was. Deze laatste strategie bestond eveneens uit audit en feedback, maar bovendien uit het verspreiden van implementatiemateriaal aan de behandelaren en folders aan patiënten, waarin de inhoud van de professionele richtlijnen in simpele taal werd uitgelegd. De gebruikte uitkomstmaat in deze trial, was de adherentie aan de indicatorenset die in hoofdstuk 3 werd ontwikkeld. De resultaten van de praktijktest uit hoofdstuk drie werd daartoe gebruikt als voormeting. Geen van beide strategieën resulteerde in een unaniem voordelig effect op de implementatie van de richtlijnen, hoewel kleine effecten werden gezien voor geïsoleerde indicatoren. Desalniettemin was het oordeel van patiënten over de folders erg positief; er werd naar aanleiding van het gebruik van de folders een verbetering van de arts-patiënt communicatie gemeld, evenals een ervaren verlaging van de drempel actief te participeren in het tot stand komen van behandelbeslissingen. Een mogelijke verklaring van het ontbreken van een effect van de geïmplementeerde strategieën, is de incomplete verspreiding van de folders onder patiënten, en een matig gebruik van het aangeboden implementatiemateriaal door de behandelaren. De meerwaarde van de maximale strategie staat bovendien ter discussie aangezien het door behandelaren meest gewaardeerde interventie onderdeel de audit en feedback was, die ook bij de minimale strategie ter beschikking werd gesteld.

Hoofdstuk 8: Algemene beschouwing

In het laatste hoofdstuk van dit proefschrift presenteren we de belangrijkste bevindingen van de studies in de voorgaande hoofdstukken. Hierbij worden de onderzoeksvragen uit hoofdstuk 1 niet alleen beantwoord, maar tevens besproken in het licht van de beschikbare en meest recente literatuur. Het blijkt dat de indicatorenset die ontwikkeld is in hoofdstuk 3, een valide instrument is om de huidige fertiliteitszorg te monitoren en verbeteren. Tegelijkertijd hebben we geleerd dat voor

elk van deze doelen afzonderlijk, wellicht een andere patiëntensteekproef of zelfs een deelselectie van indicatoren nodig is.

Voorts worden enkele methodologische overwegingen genoemd bij het interpreteren van de studieresultaten. Het gebrek aan een overtuigend effect van de gepresenteerde strategieën, sluit een toekomstige rol van patiënten in de implementatie van professionele richtlijnen echter geenszins uit. De patiëntgerichte interventie werd immers zeer positief geëvalueerd, met gerapporteerde positieve effecten op arts-patiënt communicatie en een ervaren 'empowerment' om actief mee te kunnen doen in het maken van behandelbeslissingen.

Het proberen te implementeren van een tienledig richtlijnprogramma bleek hoe dan ook een ambitieuze en tijdrovende taak. Dit is dan ook de reden dat we in dit slothoofdstuk aanbevelingen doen voor vervolgonderzoek naar patiëntgerichte implementatie van richtlijnen. Zo zou toekomstig onderzoek kunnen profiteren van een aangepast studie-ontwerp, waarin voor een meer homogene patiëntengroep wordt gekozen, of slechts een enkele, recent gepubliceerde richtlijn wordt geïmplementeerd. Om bij te dragen aan de theoretische basis van richtlijnimplementatie-onderzoek, zouden bijvoorbeeld de veranderingsbereidheid van een kliniek, team klimaat en het inzetten van lokale opinieleiders als aandachtspunten voor een vervolgstudie kunnen dienen. Tot slot worden enkele implicaties en aanbevelingen voor zowel behandelaren als beleidsmakers besproken. Zo propageren wij dat het in de nabije toekomst in te voeren Elektronisch Patiënten Dossier bij uitstek geschikt is om kwaliteitsindicatoren mee te meten, mits men daar nu reeds voorbereidende handelingen voor treft. Evenzo zal het gebruik van grotere, (inter-) nationale databases, de structurele monitoring van minder frequent voorkomende ziektebeelden of complicaties mogelijk maken. Maar bovenal zijn wij van mening dat de enorme inspanningen die heden ten dage wél worden geleverd voor het ontwikkelen van richtlijnen, gevolgd moeten worden door minstens zo grote inspanningen om die richtlijnen vervolgens te implementeren en uit die implementatie, door middel van wetenschappelijk onderzoek, lering te trekken.

Publications and presentations

Publications and presentations

Publications

- Determinants of patients' experiences and satisfaction with fertility care
  S.M. Mourad, W.L.D.M. Nelen, R.P. Akkermans, J.H.A. Vollebergh, R.P.T.M. Grol, R.P.M.G. Hermens, J.A.M. Kremer
  Fertility and Sterility Epub sept 2009.

- Information provision in fertility care: a call for improvement
  S.M. Mourad, R.P.M.G. Hermens , T. Cox-Witbraad, R.P.T.M. Grol, W.L.D.M. Nelen, J.A.M. Kremer
  Human Reproduction 2009;24(6):1420-6.

- Variation in subfertility care measured by guideline-based performance indicators
  S.M. Mourad, W.L.D.M. Nelen, R.P.M.G. Hermens , L.F. Bancsi , D.D.M. Braat, G.A.Zielhuis, R.P.T.M. Grol, J.A.M. Kremer
  Human Reproduction 2008; 23(11):2493-2500.

- Guideline based development of quality indicators for subfertility care
  S.M. Mourad, R.P.M.G. Hermens , W.L.D.M. Nelen, D.D.M. Braat, R.P.T.M. Grol, J.A.M. Kremer
  Human Reproduction 2007;22(10):2665-72.

- Monitoring reproductive health in Europe: what are the best indicators of reproductive health? A need for evidence-based quality indicators of reproductive health care.
  W.L.D.M. Nelen, R.P.M.G. Hermens, S.M. Mourad, E.C. Haagen, R.P.T.M. Grol, J.A.M. Kremer.
  Human Reproduction 2007;22:916-8.

Oral Presentations

- 'Determinants of patients' experiences and satisfaction with fertility care'
  Oral presentation at the annual meeting of the European Society for Human Reproduction and Endocrinology (ESHRE), Amsterdam, juli 2009.

- 'Improving Guideline-based information-provision in subfertility care'
  Oral presentation at the annual meeting of the Guidelines International Network (GIN), Helsinki, Finland, oktober 2008.

- 'How informed are our patients really? Guideline based information provision in subfertility care'
  Oral presentation at the annual meeting of the European Society for Human Reproduction and Endocrinology (ESHRE), Barcelona 2008, juli 2008.

- 'How informed are our patients really? Guideline based information provision in subfertility care'
  Interregionale ESHRE-bijeenkomst AMC-UMCN georganiseerd voor perifere deelnemende klinieken, Utrecht, juni 2008.

- 'De huidige OFO richtlijn: wat staat er in, en doen we wat we zeggen?'
  Regionale refereeravond voortplantingsgeneeskunde,UMC st Radboud Nijmegen, juni 2008.

- 'Kwaliteitscontrole en guidelines in de reproductieve geneeskunde'
  Voordracht op het symposium 'Kwaliteit verbeteren in reproductieve geneeskunde: ISO and beyond' Universitair Ziekenhuis st Augustinus, Antwerpen, mei 2008.

- 'Guideline based development of quality indicators for subfertility care'
  Invitational meeting of the Special Interest Group "Quality and safety" (SQUART) of the European Society for Human Reproduction and Endocrinology (ESHRE). Vision development meeting for the ESHRE guideline program', Nijmegen december 2007.

- 'Guideline based development of quality indicators for subfertility care'
  Oral presentation at the annual meeting of the Guidelines International Network (GIN), Toronto, Canada, augustus 2007.

- 'Actual and desired information provision in subfertility care'
  Oral presentation at the annual meeting of the Guidelines International Network (GIN), Toronto, Canada, augustus 2007.

- 'Guideline based development of quality indicators for subfertility care'
  Oral presentation at the annual meeting of the European Society for Human Reproduction and Endocrinology (ESHRE), Lyon, France, juli 2007.

Poster presentations

- 'Variation in clinical subfertility care measured by guideline-based quality indicators'
  35e Gynaecongres, Utrecht, juni 2009.

- 'Variation in clinical subfertility care measured by guideline-based quality indicators'
  Oral poster presentation annual conference of The International Society for Quality in Health Care (ISQua) Kopenhagen, oktober 2008.

- 'Variation in the quality of clinical subfertility care: an assessment by guideline-based quality indicators'
  Annual meeting of the European Society for Human Reproduction and Endocrinology (ESHRE), Barcelona 2008, juli 2008.

Dankwoord

Case report: A multidisciplinary approach to undo the knot

Author: S.M. Mourad, MD, a.s.a.p PhD[*]
[*]*Department of Obstetrics and Gynaecology and Scientific Institute for Quality of Healthcare, Radboud University Nijmegen Medical Centre*

Introduction

Ik wil dit proefschrift graag eindigen met het citaat aan het begin; 'Elke knoop heeft iemand om hem te ontwarren'. De SPRING-studie bleek een knoop van behoorlijk Gordiaanse omvang, en vormde voor dan ook een ware uitdaging. Zonder de hulp van vele anderen, was het nooit gelukt deze te ontwarren; iedereen die daaraan heeft meegewerkt, daarom hartelijk bedankt!

Patients

In de eerste plaats wil ik de ruim 2600 patiëntenparen bedanken die belangeloos, of liever gezegd, in het belang van een betere fertiliteitszorg, de moeite en tijd hebben genomen de uitgebreide vragenlijsten in te vullen. Zonder hulp van patiëntenvereniging 'Freya' en deze paren was het nooit gelukt zo'n goed overzicht van de huidige regionale zorg te krijgen.

The 'SPRING' study Group

Veel dank ben ik ook verschuldigd aan mijn promotores and copromotores:
Prof. dr. R.P.T.M. Grol, beste Richard, ondanks je drukke agenda en over minstens 150 man te verspreiden aandacht, wist je toch altijd precies wat er speelde in het project. Het is al vele malen eerder gezegd, maar je wist altijd met frisse blik nieuwe inzichten te berde te brengen en op praktische wijze sturing te geven aan dwalende artikelen. Bedankt voor de begeleiding de afgelopen jaren, de introductie in de wondere wereld der implementatie en met name ook je daadkracht bij de afronding van de laatste loodjes. Ik ben blij dat ik me nog tot je laatste cohort promovendi mag rekenen, want IQ healthcare zal je node missen!
Prof. dr. J.A.M. Kremer, beste Jan; ik weet nog dat ik op een koude donderdagmorgen, vers afgestudeerd, kwam solliciteren bij jou en Rosella. Ik geef het toe, Nijmegen stond niet écht hoog op mijn verlanglijstje. Maar ik was gecharmeerd van het project en hoewel ik vermoedde dat het nog weleens een hele kluif kon zijn, was ik ervan overtuigd dat het met twee open en vriendelijke begeleiders als jullie toch wel goed zou moeten komen! Daarin heb ik me gelukkig niet vergist! Ik neem in dit dankwoord even de vrijheid lekker lang van stof te zijn! Ik wil je ontzettend bedanken voor alle soepel verlopende donderdagochtend-besprekingen, je laagdrempeligheid en praktische benadering van alle beren die ik op de weg tegenkwam. Fijn dat je ons altijd van snelle reply's voorziet en altijd tijd voor je promovendi inruimt. (Zelfs toen we je aandacht moesten delen met je nieuwste iPhone ☺!)

Dr. Hermens, beste Rosella; je hebt me meer dan eens opgelapt als ik dankzij de immense omvang van de data-verzameling verzoop in de enveloppen en, de wanhoop nabij, op het spreekwoordelijke randje van de Waalbrug stond. Bedankt dan ook voor je immer goede humeur en je gave om altijd het benodigde geld of mankracht te regelen als de nood het hoogst was!

De bureau-stelling geldt zeker ook voor jou!

Dr. Nelen, beste Willianne; dank voor je oneindige precisie en het immer secuur nakijken van de zoveelste versie van mijn artikelen. Mijn beroerde typkwaliteiten en jouw altijd op volle toeren functionele spellingscheck zijn vaak met elkaar gebotst, maar het resultaat mag er zijn! Jij en Rosella zijn echt alles wat iemand zich als begeleiding in een promotie-traject wensen kan; laagdrempelig, inhoudelijk sterk, praktisch aangelegd en bovendien nog gezellig ook. Dat is lang niet voor elke promovendus weggelegd en mag dus best weleens zwart op wit staan! Ik ben daarom niet alleen blij maar ook trots dat jullie mijn co-promotoren zijn; Jan mag zich gelukkig prijzen met zo'n perfect complementair onderzoeksduo als jullie!

Prof. dr. Zielhuis, beste Gerhard; dank voor het altijd constructief en verfrissend meedenken op de projectgroepbesprekingen. Ik heb uw bijdrage altijd bijzonder waardevol gevonden!

Dr. Adang; beste Eddy; hoewel we helaas aan jouw vakgebied nooit echt zijn toegekomen, was je altijd geïnteresseerd in het project en bereid mee te denken. Dank daarvoor!

Inclusion

Speciale dank ben ik verschuldigd aan de volgende gynaecologen en fertiliteitsartsen die contactpersonen waren van de 16 deelnemende klinieken:
Peter van Dop (St Catharina Ziekenhuis, Eindhoven), Ben Willem Mol en Pettie Maas (Maxima Medisch Centrum, Veldhoven), Ton Franssen (Canisius Wilhelmina Ziekenhuis, Nijmegen), Mirjam van Rozendaal (Ziekenhuis Bernhoven, Veghel), Jos Vollebergh (Ziekenhuis Bernhoven, Oss), Ineke Minkhorst (Maasziekenhuis Pantein, Boxmeer), Carl Hamilton en Monique Brandes (Jeroen Bosch Ziekenhuis, den Bosch), Henk Ruis (behandelcentrum stichting Geertgen, Elsendorp), Guido Muijsers (Rivierenland Ziekenhuis, Tiel), Jules Schagen van Leeuwen en Tessa Cox (St Antonius Ziekenhuis, Nieuwegein), Bert van Dam en Thecla van de Kamp (Gelre Ziekenhuis, Apeldoorn), Eduard Scheenjes en Caroline Leendertz (Ziekenhuis Gelderse Vallei, Ede), Alexander Schmoutziguer en László Bancsi (Ziekenhuis Rijnstate, Arnhem), Vincent Blom (Streekziekenhuis Zevenaar), Eveline Tepe, (Slingeland Ziekenhuis, Doetinchem).
En uiteraard even zovele secretariaten, poli- en archiefmedewerkers en fertiliteits-teams ter plaatse voor hulp bij dossiers zoeken, dataverzameling en distributie van vragenlijsten en patiëntenfolders.

Externe co-auteurs, László Bancsi, Jos Vollebergh en Tessa Cox, bedankt voor jullie bijdrage en correcties bij de betreffende artikelen.

Material and Methods

Met stip op nummer één: Janine, zonder jouw oneindige hulp en inzet was het met de hele inclusie, dataverzameling en die kilometers syntax nooit gelukt! Ondanks het feit dat ik altijd het onmogelijke in veel te korte tijd wilde, bleef je altijd even precies. Zie figuur 1.
(PS : …check jij die syntax nog even ?)

Figuur 1: SPSS syntax dankbetuiging

```
COMPUTE dank =99.
DO IF object= 'proefschrift' AND uniekid_1=
'selma' AND uniekid_2= 'janine'  Dank =4.
 EXECUTE.
VALUE LABEL dank '0' niet aanwezig '1'
neutraal '2' gemiddeld '3' groot 4 'duizendmaal'
99 'ongeldig'.
Variable label uniekid_1 ' degene die dank
betuigt'  uniekid_2 'degene die dank ontvangt'
MISSING VALUES dank '(7 thru 9).
```

Reinier, je hebt de zeldzame gift statistiek logisch te laten klinken en het begrijpelijk uit te leggen. Dank voor al je hulp bij de multilevelanalyses!

Het gehele secretariaat IQ, en met name Jeannette van 'sectie 2' voor allerhande hulp bij vooral scannen en verifiëren van oneindig veel Teleform-vragenlijsten en spoed-verzendingen.
Jolanda, heel erg bedankt voor je hulp bij het gereed maken van het manuscript en het mij achter de broek zitten toen de tijd begon te dringen!

Data collection:

Onderzoeksondersteuning en onmisbare praktische hulp kreeg ik van Annelies, Sabine en Marc: bedankt! Studenten Evelyn, Anke, gezondheidswetenschapper Petra hielpen bij de voormeting en studenten Jasper en Monica droegen behoorlijk wat steentjes bij aan de nameting, naast nog enkele bijklussende 'verzend'-studenten.
Samen worstelden we ons door ruim 2600 dossiers, en er werden in totaal ruim 6000 vragenlijsten en reminders verstuurd en dus even zovele brieven gevouwen, enveloppen gevuld, bestickerd en dichtgeplakt. Dat we er al met al maar een pijnlijke 50 stuks zonder inhoud hebben verzonden, is dus hoogstwaarschijnlijk niet significant ☺!
Marc, Mark F en Gilbert, bedankt voor frequente, acute ICT-helpdeskerij.

Setting

Co-workers

Alle collega's van de Pijler Voortplantingsgeneeskunde van het UMCN: Daarbij noem ik natuurlijk de VPG gynaecologen Jan, Ina, Wim, Didi en Hans voor de leerzame tijd tijdens mijn eerste schreden in de kliniek; de immer goedgehumeurde en goedgeklede dames van het Secretariaat voor hulp bij praktische zaken; analisten en embryologen van het lab waar ik terecht kon voor al m'n 'zaadvragen' of het bietsen van koekjes, alle verpleegkundigen en NP'ers Leny en Jolieneke voor de fijne samenwerking en eigenlijk altijd gezellige weekenddiensten (zeker toen het koffie-apparaat eenmaal daar was!). Ik heb ruim 4 jaar met veel plezier bij én met jullie gewerkt, de Freya-award is op zijn plek!
En last, but certainly not leas,t collega IVF-artsen Bea, Gwendolyn, Esther, Anika en Joyce. Jullie waren fijne collega's! Dames, zorg dat de VPG minstens een doctoraat-hattrick scoort dit jaar!

Alle collega's uit de kantoortuin[1] en LAT-relatie Emmy, die bezig of net klaar zijn met hun eigen promotie. Hoewel ik nooit vaker in de tuin heb gezeten dan het afgelopen jaar aan zaterdagen en zondagen, was daar altijd minstens nog één kantoortuiner keiharde wetenschap aan het bedrijven[2]. Maar dankzij de frequente publiceren-is-trakteren-taart was er vaak genoeg

reden om over de vloer te komen en ook de onderzoekers-weekendjes waren ronduit memorabel. Succes met het afronden van jullie promoties, ik hoop dat we snel weer allemaal collega's zijn!!

Collega's IQ, teveel om op te noemen, dank voor de hulp, leerzame tijd vol printerperikelen, afdelingsuitjes, refereeractiviteiten. Ik noem speciaal de IQ-skitrip en gezamenlijk afzien bij de zevenheuvelenloop. De IQ-kelderbewoners bedankt voor de gezelligheid toen we naar -1 'verbannen' werden, en specifiek Geert voor het vele malen aanhoren van onderzoeksgerelateerde stress boven het Senseo-apparaat danwel een Aesculaafbiertje met Lucy en Marije. Lucy, ondanks enige topografische dwalingen zouden we het bezoeken van watervallen (Howcome, Niagara ain't no Kitchener?), vulkanen (no puedo ver El Arenal?), bergtoppen (porqué ninguna agua caliente?) en besneeuwde hellingen (hast du ein Eisbär gesehen?) zeker niet voor gezien houden;-)

Collega's Rijnstate: Na een leuke tijd in het Radboud vroeg ik me af wat me in Arnhem te wachten stond. Gelukkig was die zorg onterecht! Maatschap gynaecologen, bedankt voor de warme ontvangst, leerzame tijd en goede werksfeer die er in het Rijnstate heerst. Collegae A(N)IOS[3] : we zijn een leuke ploeg! Bedankt voor jullie aanhoudende interesse in mijn chronisch 'nu echt bijna klaar'-levenswerk: snel weer borrelen! Alle klinisch verloskundigen, poli-medewerkers, verpleegkundigen, kraamverzorgsters, bedankt voor de fijne samenwerking en geduld met dit groentje! Han, volgend jaar weer gewoon drie keer naar de sneeuw ☺!

Various contributors

Lieve clubgenootjes Absoluut[4] het moet maar weer eens gezegd dat we een bijzonder leuk stel zijn! Wat een feest dat we elkaar ondanks geografische spreiding na 11 jaar nog leuk genoeg vinden om regelmatig uitbundige dates te maken. Ik hoop dat er nog maar veel dr's, weekendjes weg, hilarische lustrumfotoboeken en 'over-the-top' etentjes mogen volgen!

Klaar, Tamaar en Diek, succes met jullie proefschriften! Ier en Klaar, ik ben blij dat de onovertroffen alfabetische volgorde ons in 1997 in hetzelfde introductiegroepje deed belanden! Snel weer samen fietsen en feesten!

Oud-ploeggenoten Hille, Mirthe, coach Eva en andere Njordvriendjes en vriendinnetjes. De 'Head of the River Amstel' bij windkracht 6 tegen is niets vergeleken met een promotietraject…Maar na de finish is het nog altijd goed bieren, dus trek de agenda's maar!

NMHC dames 11 en 8, het is nog steeds lekker rondjes rennen met jullie! Lon, superbedankt voor je grafische hulp bij het ontwerp van de mooie kaft en tussenbladen!

Dames van de Nieuwe Rijn 80, dát waren tijden, wat is het toch altijd leuk met jullie! Hopelijk tot snel op een oud-HG-event!

Location

De Legendarische kamer 1.04 op het WOK later KWAZO tegenwoordig IQ, was de reden dat ik de kantoortuin links liet liggen. Hier werd niet alleen hardcore wetenschap bedreven, maar ook toekomst- en vakantieplannen bedacht, afwijzingen en publicaties

gevierd, afterlunch gedipped, ge-brand-een-kaarsje-punt-NL-ed en vanaf dag één 'stellingen voor als het ooit zover is' verzonnen.

Lieve Marije, Bosschie, 1.04-roomie nummer 1; zonder de andere dames nimfen teniet te doen, jij had natuurlijk aan mijn zijde moeten staan…. M'n liefste 'Nijmeegse' vriendinnetje, die zo verdomd goed haar best doet zover mogelijk uit de buurt van Nijmegen te blijven! Gelukkig heb je met een gloedvolle nieuwe carrière (én een riant logeerbed) in Melbourne, de enige goede reden om te ontbreken. Dat gezegd hebbende, moet je wel beseffen dat het je nog heel wat bierella's en vrijdagavond-kipsaté gaat kosten om dít weer goed te maken! Misjoetoe;-)

Paranimfen

Dan natuurlijk mijn allerliefste paranimfen die mij met raad, daad en hun sprankelende verschijning zullen bijstaan:

Lieve Esther, 1.04-roomie nummer 2; Af en toe deelden we ook 4 muren op spannender locaties zoals Lyon en Barcelona, of op de lange VPG-woensdagmiddagen 'poli met z'n twee op min één'. In elk geval was het met jou altijd leuk! Maarre…had je ze nou niet even kunnen overtuigen in dit project alleen die IUI richtlijn te implementeren;-)? Super dat je mijn paranimf wil zijn! Veel succes bij je eigen 'laatste loodjes' en afronding van je proefschrift, almost there!

Lieve Jakkes, inmiddels zijn we zo ongeveer langer samen op vakantie geweest dan dat we eigenlijk samen gewoond hebben; en toch ben je m'n favoriete NR80-huisgenoot! Zet hem op bij je eigen promotie-onderzoek en blijf vooral de wereld zien in rebussen en expressieve mandarijntjes;-). Oja, en nu wil Snel snel weer een weekendje met je weg!

Acknowledgements

Lieve papa en mama, nummertje drie is ook eindelijk drs-*af*. Daar mogen jullie nu ook eens wat credits voor krijgen, want die goede genen komen heus ergens vandaan! Daarbij hebben jullie ons de brede interesse voor cultuur en natuur bijgebracht en ons, hoe eigenwijs ook, nooit een strobreed in de weg gelegd als het ging om studie of talrijke buitenschoolse activiteiten. Pap, nu drie doctors en twee dokters onder 1 dak, met bijbehorende ego's en praatjes; je moet er maar tegen kunnen! Gelukkig is je humeur onverwoestbaar en je geduld oneindig. Ik denk dat ik nu wel weer tijd heb voor een gezamenlijke Sudoku! Mam, '*Three down, one more to go* '; Nu is het *mijn* beurt om te kunnen zeuren wanneer die promotie, als kroon op je carrière, nou eindelijk eens afkomt ☺, zet hem op!

Daniël en Maurice, jullie 'HBO'-zusje staat weer op gelijke voet. En tja, het dankwoord is langer dan twee pagina geworden, who cares!

En nu; tijd voor vakantie, een boek en een andere knoop!

References

[1] Gwendolyn, Willianne, Wouter, Joyce, Suzan, Refika, Roosmarie, Bea, Eva, Anne, Angèle, Thijs, Suzan, Hedwig, Irene, Ineke, Charlotte (2x), Sabine, Inge, Annemijn, Elvira, Marian, Ralph, Channa, Linda, Arno, Esther, Anika

[2] Irene, alle weekends@work waren gezellig!

[3] Suzanne, Jantien, Marieke J en Marieke S, Tjitske, Anne, Hanneke, Sanne, Moniek, Marjolijn, Dorus, Jet, Josine, Colette, Eva, Kim

[4] Klara, Iris, Steph, Steef, Dieke, Isa,Tamara, Ilena, Vanessa, Jennifer, Ingrid (2x), Cecile

Curriculum Vitae

Curriculum Vitae

Selma Mourad werd geboren op 1 maart 1980 te Leiden en groeide op in het nabijgelegen Leiderdorp. Bij gebrek aan duidelijke toekomstplannen volgde zij op het Stedelijk Gymnasium Leiden zoveel mogelijk talen alsook exacte vakken, waarop zij in 1997 het VWO-diploma behaalde. Direct daarna werd aangevangen met de studie Geneeskunde aan de Universiteit Leiden. Tijdens haar studententijd woonde zij 6 jaar met veel plezier met 20 andere dames, 1 gezamenlijke TV en kat Harrie op de Nieuwe Rijn 80. Daarnaast werd er gelezen, gehockeyd, hardgelopen, geborreld, geroeid en vele malen het Rondje Ringvaart gefietst. In het seizoen 1999-2000 nam zij als ab-actis plaats in het bestuur van de Koninklijke Studenten Roei Vereniging 'Njord', waarna zij nog twee jaar actief was als wedstrijdroeier, en tot op heden als Bestuurslid van oudleden-vereniging 'Oud Njord'.

In 2001 vertrok Selma voor 4 maanden in het kader van een wetenschappelijke stage naar de Verenigde Staten, waar zij in Portland (Oregon) bij het Rosenfeld Lab, verbonden aan het Doernbecher's Children's Hospital van de Oregon Health and Science University (OHSU), onderzoek deed naar de rol van Insuline Like Growthfactors en bijbehorende Binding Proteins op geprogrammeerde celdood (begeleiders dr. Junko Tsubaki, dr. Vivian Hwa, Prof dr. Ron Rosenfeld). Vele Western, Northern blot's en de onfortuinlijke dood van met liefde verzorgde celkweken verder, was het duidelijk dat er, ondanks de schoonheid van the Pacific Northwest, zeker geen lab-carrière voor haar in het verschiet lag. Vervolgens begon Selma aan de co-assistentschappen, en verbleef onder andere 3 maanden in Beirut, Libanon, voor het volgen van de coschappen Dermatologie en Keel-, Neus- en Oorheelkunde aan de American University of Beirut Medical Centre (AUBMC). Eind 2004 behaalde zij het artsexamen en werd direct aangesteld in het UMC st Radboud als artsonderzoeker op de door ZonMw gesubsidieerde 'SPRING-studie', geïnitieerd door het Scientific Institute for Quality of Healthcare (IQ healthcare, voorheen afdeling Kwaliteit van Zorg) en de afdeling Verloskunde en Gynaecologie van het UMCN. Een vruchtbare samenwerking tussen deze twee afdelingen in de personen van enerzijds Prof dr. Richard Grol en dr. Rosella Hermens en anderzijds Prof dr. Jan Kremer en dr. Willianne Nelen, resulteerde in de studies beschreven in dit proefschrift. Naast het doen van onderzoek, was Selma bovendien enkele dagen per week werkzaam als IVF-arts bij de pijler Voortplantingsgeneeskunde. In mei 2009 begon zij als ANIOS te werken op de afdeling Verloskunde en Gynaecologie van het Rijnstate Ziekenhuis (opleider dr. F.P. Dijkhuizen). In juli 2010 zal zij officieel met de opleiding tot Gynaecoloog starten in het Nijmeegse cluster, en wel in het Ziekenhuis Gelderse Vallei te Ede (opleiders Prof. dr. D.D. Braat en dr. E. Scheenjes).